INTRODUCTION A LA LOGIQUE PERTINENTE

François Rivenc

Introduction à la logique pertinente

François Rivenc

ἠνίδε κου κόρακεσ τεγέων ἔπι κοῖα συνῆπται κρώζουσιν καὶ κῶσ αὖθι γενησόμεθα.*

* Voilà comment les corbeaux sur les toits croassent quelles sont les implications, et comment nous vivrons.

Callimaque, cité par Sextus Empiricus,
Adv. Math. i. 309.

REMERCIEMENTS

Toute ma reconnaissance va à mes étudiants de la Maîtrise de logique de Paris 1, durant les deux années universitaires 2002-2003 et 2003-2004 : ils ont bravement accepté de pénétrer avec moi dans la jungle des logiques pertinentes, et souvent leurs questions m'ont aidé à aller de l'avant ! Elle va également à Anne-Gabrielle Wersinger, maître de conférences à l'UFR de philosophie de Paris I, et spécialiste de logique antique, dont la présence régulière à mon cours nous aida à faire le lien entre les discussions grecques et les questions actuelles ; elle nous apporta le charme et la subtilité d'une tradition quelque peu oubliée.

Dans un registre plus intime, je remercie Amaro de Villanova et Antoine Fontaine : ils sauront pourquoi. Et aussi Sylvie Kipen-Rivenc, *for·ever.*

Mes remerciements s'adressent aussi à Dominique Lecourt, qui soutint dès le début le projet de cet ouvrage, et l'accepta avec enthousiasme dans sa collection.

Enfin je ne sais comment remercier Bernard Victorri, qui consacra de longues heures à numériser une version de cet ouvrage, lui rendant par là une vie menacée par sa malheureuse disparition en tant qu'objet physique !

Sommaire

Préface

> « La voie suivie par Anderson et Belnap dans *Entailment*, afin d'atteindre leur système préféré **E**, ouvrit la voie à une grande variété, une pléthore de systèmes - ouvrit la boîte de Pandore. »
>
> Norman et Sylvan, 1989.

On l'a dit et redit : pour nombre de philosophes français, ·la logique est encore l'objet d'une ambivalente fascination. L'idée est sans doute que la logique est le domaine de la rigueur absolue, des preuves impeccables : comment discuter ce qui légifère sur les conséquences nécessaires de vérités elles-mêmes nécessaires ? D'où la tentation d'utiliser de manière parfois absurde ses « résultats », dès lors qu'ils paraissent profonds, et de nature à conforter une vision philosophique. D'où aussi la tentation inverse d'y voir une prison pour la liberté de l'imagination philosophique ! Je n'insisterai pas.

Puisqu'il est salutaire, cependant, de commencer par reconnaître ses torts, il faut avouer que la manière dont la logique est enseignée dans les départements de philosophie porte une part de responsabilité dans cet état de choses. À l'opposé de l'esprit d'exploration et de remise en question, qui est censé être le lot de l'enseignement philosophique (je dis bien censé*), la logique est présentée de manière beaucoup trop dogmatique*[1]*. À côté de la philosophie - mais en fait* hors *d'elle -, ce sont d'emblée les rudiments de la logique moderne* classique *qu'on enseigne, ceux de la logique mathématique, telle qu'elle a été progressivement mise en forme, disons, des années 1920-1930 au début des années 1950. Du point de vue des techniques formelles, il faut commencer par là, c'est évident, en raison de la simplicité, de l'élégance ; et du côté « carré »* de *cette logique. Mais du point de vue conceptuel, ne faudrait-il pas au moins la* situer *? Clarifier les choix qui ont présidé à sa construction, et qui doivent être évalués ? Comment réagir, par exemple, à la foisonnante pluralité des systèmes logiques d'aujourd'hui ? Des philosophes comme Quine ont présenté des arguments à l'appui de la thèse que l'idée de la logique était définitivement incarnée par le système classique du premier ordre. En gros, l'argument est que l'essentiel des capacités*

1. Épinglant ce dogmatisme, Graham Priest ajoutait en 1979 : « Je pense que quiconque parmi nous a enseigné la logique classique sera d'accord avec cette observation, et je suis sûr qu'elle frappera une corde sensible chez la plupart des enseignants de logique» (« Classical Logic *aufgehoben* », *in* Priest, Routley et Norman, 1989).

inférentielles de la quantification y est formalisable. Quoi que l'on pense de ce type d'arguments, il y a déjà un progrès d'accompli : la discussion est ouverte, serait-ce sous la forme fruste : « Où est la vraie logique ? » (je ne pense pas que ce soit là exactement la bonne question*).*

Stephen Read, dans son excellent livre Thinking about Logic *(Read 1995), fait remarquer que l'idée que la logique, invention après tout humaine, incarne la rigueur absolue, est fondée sur une petite confusion. Certes, si B est conséquence logique de A, c'est peut-être là un état de choses nécessaire, auquel nul ne peut rien changer ; de même, qu'une proposition soit logiquement vraie ne dépend pas de notre bon vouloir. Mais il ne s'ensuit nullement que nos affirmations concernant le fait que B soit conséquence de A soient également indiscutables ; encore moins qu'une théorie complexe, qui tente d'expliquer* pourquoi *nous devons apprécier ainsi le lien entre A et B, ne puisse se tromper. De même que des hypothèses scientifiques ayant un pouvoir explicatif réel concernant un champ de phénomènes peuvent se heurter à de nouveaux phénomènes résistants, il se peut que la théorie logique classique, appropriée à certaines formes d'inférence, finisse par donner une image déformée d'inférences que nous reconnaissons spontanément comme valides, ou au contraire invalides, ou encore douteuses. Il est donc naturel qu'il y ait - et ce depuis la naissance de la logique moderne, voir la logique intuitionniste -, des logiques non classiques, des logiques dites « déviantes ». A lui seul, cet état de choses n'est pas de nature à dissoudre dans le pluralisme l'idée de la* vraie et unique logique *; il se pourrait qu'il y ait quelque chose comme la vraie logique, et des prétendants nombreux rivalisant dans leurs prétentions à la représenter (pour dissoudre l'idée de la logique, il faut d'autres arguments, proprement philosophiques). Comme on le verra, j'irai plus loin que Read sur cette pente : je ne pense pas qu'il existe un concept unique et assuré de conséquence logique, que nous le saisissions correctement ou non à travers tel ou tel système.*

Aristote introduit les Premiers Analytiques par la définition selon laquelle « le syllogisme *(on pourrait dire : une inférence en général) est un discours dans lequel, certaines choses étant posées, quelque chose d'autre que ce qui a été posé résulte nécessairement du fait que ces choses soient telles ». Au cœur - ou à l'horizon, comme on voudra -, de la théorie logique, réside depuis sa naissance le concept de conséquence. Un système de logique déterminé a pour tâche de construire un concept de conséquence clairement défini, par divers moyens, tantôt dits « syntaxiques », tantôt dits « sémantiques », le plus souvent les deux à la fois (et c'est une grande satisfaction que de pouvoir montrer que les deux définitions coïncident extensionnellement, ce que donne un théorème de complétude ; car cela renforce la conviction qu'on a vraiment défini une notion stable, qu'on peut retrouver à travers différents points de vue). Parlons pour l'instant d' « implication logique » à propos du concept de conséquence formellement défini en relation avec un système de logique. La propriété de complétude peut se formuler ainsi : la notion de* déductibilité *(qu'est-ce qu'on peut déduire de telles prémisses ?) recouvre exactement la notion d'implication logique ; quand on déduit conformément aux règles du système, on est sûr qu'on procède selon des relations d'implications. Et réciproquement, toute relation d'implication logique peut être représentée par une déduction formelle. La logique classique du premier ordre, qui est complète en ce sens, retire de ce fait, bien naturellement, une grande force de séduction.*

La philosophie de la logique commence avec la question beaucoup plus difficile (conceptuellement, s'entend), de ce que l'on pourrait appeler l'adéquation *de la complétude. Le concept d'implication, relatif à un système, a-t-il un champ assez vaste pour couvrir tous les cas d'inférences justifiées par une authentique relation de conséquence (« complétude » de la complétude) ? Ou bien, au contraire, n'entraîne-t-il pas des .diagnostics erronés, légitimant comme valides des inférences surprenantes pour un jugement non déformé par l'apprentissage intensif du système ? Si tel est le cas, on peut sérieusement douter que le concept théorique soit une bonne explication, au sens de Carnap, du concept... comment faut-il dire : ordinaire ? intuitif ? préthéorique ? Mais qu'est-ce qu'un « concept intuitif » ? Faut-il se fier à cette expression usuelle, mais peut-être trop commode ?*

Un exemple très simple, pour commencer, aidera à saisir le point (c'est une forme de l'un des fameux « paradoxes » de l'implication matérielle). On peut sans doute admettre sans trop de difficulté que des deux prémisses, en fait incohérentes :

> *Le poids d'un corps est un nombre réel positif,*
> *Une substance d'un poids négatif est libérée lors de la combustion d'un corps,*

s'ensuive la conclusion (vraie, c'est un fait d'observation un peu soigneuse) :

> *Lors de la combustion, le poids total du corps brûlé augmente.*

La logique classique ne conteste pas la correction de cette inférence, au contraire. Elle l'explique même en faisant remarquer que de prémisses contradictoires n'importe quoi s'ensuit. Pourquoi ? Selon elle, ce qui caractérise avant tout la relation d'implication, c'est son caractère de préservation de la vérité. *Dire que A implique B, c'est dire que si A est vrai, B l'est aussi. En d'autres termes, qu'il n'est pas possible que A soit vrai et B faux. On conclut de là que A implique B, si et seulement si on n'a pas A vrai et B faux (moyennant un certain traitement de la modalité exprimée par» il n'est pas possible que »). Or si A est contradictoire, il n'est pas vrai ; donc quel que soit B, on n'a pas A vrai et B faux. Donc A contradictoire implique B, pour n'importe quelle proposition B. D'où le principe :* Ex falso quodlibet sequitur.

Si cette explication est adéquate, alors des deux mêmes prémisses, Priestley était aussi bien autorisé à conclure correctement :

> *Certains corps sont de bons conducteurs de l'électricité,*

(Priestley, paraît-il, a travaillé sur les phénomènes électriques). Mais il est fort douteux que d'une théorie absurde de la combinaison d'un corps avec l'oxygène lors de la combustion, s'ensuive, en un sens communément admis du terme, n'importe quelle conséquence concernant la conductibilité électrique. Et on fera remarquer que les prémisses, contradictoires ou pas, sont sans pertinence *pour la conclusion (bien sûr, on peut contourner l'argument, en disant que ce qui nous trouble, c'est que l'inférence n'est pas intéressante, justement parce que d'une théorie contradictoire on peut déduire n'importe quoi ; mais que cela n'a rien à voir avec sa correction !). Quoi qu'il en soit, la moralité que tire de l'exemple la* logique pertinente *est la suivante : des considérations plus fortes que la seule préservation de la vérité, la prise en compte de relations de contenu doivent présider à la construction d'un concept authentique de conséquence*

logique. La logique classique sanctionne comme valides des inférences, qui, pour un regard non prévenu, ne le sont pas.

Cet ouvrage n'est pas une revue des logiques dites» non classiques », « déviantes », ou « philosophiques », mais seulement une étude consacrée à l'une d'entre elles, la logique pertinente (traduction du terme anglais « Relevant Logic ») : la prise en compte de la pertinence des prémisses pour la conclusion vient d'être évoquée au paragraphe précédent. Encore faut-il préciser. On parle plus volontiers des Relevant Logics, au pluriel, en entendant par là les différents et nombreux systèmes de logique pertinente, comme si la logique pertinente n'était elle-même qu'une idée régulatrice, susceptible de s'incarner dans différents systèmes : les systèmes **B**, **T**, **R**, **E**, **RM**, pour n'en citer que quelques-uns parmi les plus familiers. J'ai renoncé à passer en revue ces systèmes, dans la pensée qu'une revue est facilement ennuyeuse : le lecteur peut avec raison être découragé devant une boîte à outils si abondante, outils dont il n'est pas sûr qu'ils puissent jamais servir. Au centre de l'attention figure donc, tout au long de l'ouvrage, le seul mais central système **R** de l'implication pertinente, sous ses diverses présentations : axiomatique, en Déduction naturelle, en Calcul des séquents. Les remarques éventuelles concernant les autres systèmes pertinents ne seront que de brèves échappées.

Si je devais justifier, autrement que par mes préférences personnelles, l'abord du concept de conséquence par le biais de la logique pertinente, je dirais ceci. D'une part, la logique pertinente, parmi tout ce qui figure sur le marché comme logique non classique, est incontestablement une théorie logique pure, ce qui n'est pas toujours le cas d'autres théories qui se présentent aussi comme des logiques, logiques de ceci ou de cela, ce qui devrait suffire à signaler leur caractère de logique appliquée. On parle ainsi volontiers des logiques épistémiques, doxastiques, etc. Il s'agit ici de formaliser des notions, celle de savoir par exemple, en proposant des axiomes plausibles destinés à rendre compte de leur comportement, en particulier dans des contextes inférentiels. On reconnaît ainsi que le fait qu'un sujet A sait que p, implique que p. On peut se demander également si A sait que p entraîne que A sait qu'il sait que p (Alain répondait oui : « Savoir, c'est savoir qu'on sait »). La notion de savoir est-elle incontestablement logique ? Même en faisant la part du caractère relativement arbitraire des classifications, on pourrait aussi bien soutenir qu'il s'agit de rechercher un système axiomatique reflétant les traits essentiels d'un concept épistémologique. Bien sûr, comme dans toute recherche axiomatique, de la logique est investie, ou appliquée ; mais la question ne porte pas spécialement sur la nature de la logique mise en œuvre. Il en est essentiellement de même pour les « logiques conditionnelles », qui sont en fait des théories sémantiques des conditions de vérité des conditionnels ordinaires (des expressions en « si-alors ») : les inférences sanctionnées ou invalidées n'ont au fond qu'une valeur de test pour la théorie proposée.

La logique pertinente, au contraire, a pour objet essentiel le concept même de conséquence, concept « logique » si quelque chose l'est, ou du moins un certain concept de conséquence. En fait, le nom est quelque peu trompeur, bien qu'il le soit moins que « logique de la pertinence » (« Relevance Logic »), s'il suggère qu'il s'agirait d'étudier une notion globale, la pertinence, qu'il s'agisse de la pertinence d'une argumentation, d'une assertion dans un contexte, d'une intervention dans une discussion, etc. Il vaudrait mieux parler de « logique de la conséquence pertinente », et c'est bien ainsi *que l'entendaient ses véritables fondateurs, Alan Ross Anderson et Nuel Belnap. Savoir si la conséquence pertinente est la vraie notion de conséquence, si la logique pertinente est en fait la* vraie *logique, est une autre question : faut-il être* pertinentiste *? Ma réponse ultime sera :* non. *Mais en attendant, je suis disposé à soutenir : que la logique pertinente a pour elle d'épouser souvent plus fidèlement que la logique classique les sentiments d'adhésion et de rejet que nous pouvons avoir devant certains schémas ou principes d'inférence ; de justifier, au moins partiellement, un ensemble assez stable de théorèmes ou de règles d'inférence qui possède une véritable cohérence : précisément le système **R**. Pour une théorie, adéquation, force explicative et cohérence sont évidemment de grandes qualités.*

Mais il est évident que la logique pertinente pose de nombreux problèmes, dont la mise à jour jalonne, pour ainsi dire, son histoire. Le rejet, par exemple, du principe d'inférence connu comme le « Syllogisme disjonctif » (SD) peut paraître insoutenable. En fait, comme on le verra, l'analyse de ce point mène droit au cœur des questions les plus cruciales que la réflexion sur la logique pertinente permet de soulever : la nature de la négation, les voies et moyens de sa représentation, et pour finir la place et la portée de la sémantique relativement aux présentations syntaxiques de la logique. Je mets au crédit de la logique pertinente d'obliger le logicien à prendre, ou à reprendre, ces questions à bras le corps.

On touche .là à la seconde raison que je pourrais invoquer du privilège accordé à la logique pertinente dans cet essai. L'étude des questions qu'elle soulève mène naturellement à ce que j'appellerai le point de vue structural, *qui peut se recommander, à mon sens, des avantages suivants : il éclaire d'un jour nouveau la question de la négation (y a-t-il réellement quelque chose comme une « négation pertinente » ? de ce point de vue, la réponse est encore :* non*). Il remet la sémantique formelle à sa juste place, de n'être qu'une représentation ou une construction* ad hoc *de concepts dont la légitimité et .la justification viennent d'ailleurs, en dépit de l'impression contraire que donne la sémantique appropriée à la logique classique. Enfin et surtout, il propose une solution élégante et fondée à la question de l'unité de la logique, au-delà de la pluralité des systèmes logiques qui trouvent ainsi leur place dans un cadre conceptuel unifié. La position que je propose n'est donc pas « pertinentiste », parce qu'elle admet le pluralisme des systèmes ; mais il ne s'agit pas d'un pluralisme débridé, « gauchiste », dans la mesure précisément où la notion de l'unité de la logique y trouve son compte : Quelques mots préliminaires sur ce point de vue structural permettront au lecteur de savoir où il va.*

Il est usuel, à présent (selon Kosta Došen, depuis un séminaire tenu à Tübingen en 1990), de ranger la logique pertinente, au sein des logiques non classiques, parmi les logiques dites « substructurales », à côté par exemple de la logique intuitionniste, de la logique linéaire, etc. Cette caractérisation trouve son origine dans des travaux de Gentzen, le Calcul des séquents. Parmi les règles de son système, règles qui autorisent le passage d'une inférence à une autre (ou figures de déduction), Gentzen distinguait les règles structurales *des règles opératoires qui elles seules concernent les connecteurs logiques :* et, *ou, non, si-alors*[1]*. Les règles structurales concernent au contraire ce*

---

1. Je reprends l'adjectif « structural » à la traduction française par Feys et Ladrière des « Untersuchungen über das logische Schliessen » de Gentzen (Gentzen, 1934).

qu'on appelle volontiers aujourd'hui la gestion des prémisses : si par exemple A et B impliquent une conclusion C, on peut en inférer que B et A impliquent C (règle de Permutation). Une autre règle structurale est la règle d'Affaiblissement : si A implique C, on peut rajouter une prémisse quelconque B ; les deux prémisses A et B impliquent également C (propriété parfois dite de « monotonie » de la relation de conséquence classique).

La règle d'Affaiblissement (Genzten utilisait le terme allemand « Verdünnung ») n'a rien d'absolument évident. Gentzen a éprouvé, semble-t-il, le besoin de la justifier, précisément en invoquant la difficulté qu'il y aurait à la restreindre aux cas où il y aurait un « lien de dépendance » entre la nouvelle hypothèse et les autres propositions. Il écrit ainsi à propos de cette règle :

> *« À première vue, elle peut sembler étrange. Pourtant, par exemple, si une proposition est vraie, nous sommes contraints d'admettre que dans ce cas elle vaut aussi sur la base d'une assomption arbitraire (si nous voulions stipuler que cela ne peut être asserté que dans les cas où une dépendance factuelle [tatsächliche Abhängigkeit] existe, des difficultés considérables surgiraient, en raison de l'existence de preuves où seul un usage apparent d'une assomption est fait) » (Gentzen, 1936, cité par Kosta Došen, in Schroeder-Heister et Došen, 1993).*

*Les logiques qui, dans leur présentation en Calcul des séquents, rejettent ou restreignent certaines des règles structurales de Gentzen sont dites pour cette raison « substructurales ». Or c'est un fait remarquable qu'en faisant un usage intensif de la notion de collections structurées de .formules (prémisses, et éventuellement conclusions multiples), l'ajout ou le rejet de telle règle structurale donne des systèmes bien identifiés, ou équivalents à des systèmes bien connus par ailleurs, les règles opératoires pour les connecteurs restant inchangées de système à système. En particulier, l'absence de la règle d'Affaiblissement donne précisément le système **R** de logique pertinente, sous sa .forme « gentzenisée ». Comme le maintien des mêmes règles opératoires permet de penser que la signification des connecteurs n'est pas affectée lors du passage d'un système à l'autre, on a là un argument puissant contre l'idée que « changer de logique, c'est changer de sujet », parce qu'on ne parlerait plus des mêmes constantes logiques. Ce point est en particulier crucial pour la négation, comme on l'a suggéré plus haut. De plus, cette» théorie générale des relations de conséquence structurée » (Gabbay,* in *Schroeder-Heister et Došen, 1993) est susceptible de fournir le cadre conceptuel unifié où continue à vivre l'idée de l'unité de la logique.*

*De quelque manière, au cours de cet ouvrage, l'étude de la logique pertinente se révèle être une bonne voie d'accès (et, du point de vue de la justification conceptuelle, la meilleure que je connaisse), à ce point de vue structural. Au bout du compte, nous aurons abandonné l'idée d'*une *relation de conséquence logique, au profit d'une vision plus flexible d'une famille de telles relations. Il est temps d'expliciter ce que j'ai déjà laissé entendre ici et là, à titre de justification de ce pluralisme modéré. Avons-nous donc un concept « intuitif », préthéorique, de ce qu'est une inférence valide ? Carnap, quand il s'occupait de justifier la procédure de l'explication, parlait volontiers de remplacer un concept* vague, *repérable dans l'usage, par un concept précis, formellement construit. Qu'est-ce qu'un concept vague ? Je serais plutôt ici du côté de Frege, qui tenait qu'un concept vague n'était pas du tout un concept (mais une contradiction dans les termes). S'il y a du vague, si les frontières sont mal délimitées, si* *l'usage délivre des verdicts incertains, c'est que nous manquons d'un concept, ou plutôt qu'il n'y a rien de tel : d'où l'intérêt, justement, des .formalisations. Je ne crois donc pas qu'il y ait* un *concept naturel, intuitif, ou ce qu'on voudra, de conséquence logique, qui puisse soutenir l'idée qu'il y aurait une vraie ·logique, celle qui capturerait ce concept. On admet communément que les théories scientifiques sont sous-déterminées par la totalité des données disponibles, même en principe. Pourquoi ne pas admettre, pareillement, que les procédures de .formalisation conceptuelle sont sous-déterminées par la totalité des données, qui sont au moins aussi complexes, lacunaires et indécises, que dans le domaine des phénomènes naturels ? Les données sont ici d'ordre linguistique : que disons-nous, et qu'inférons-nous de ce qui est dit ?, mais aussi en partie d'ordre réflexif, sinon déjà théorique : les appréciations portées sur la validité d'une inférence sont souvent instruites par des jugements antérieurs, fruits d'examens critiques transmis par la tradition. Les données peuvent être aussi comme filtrées par des principes généraux : c'est sans doute le cas en ce qui concerne le principe d'extensionnalité, l'idée d'une logique extensionnelle. Face à cette réelle complexité, mieux vaut sans doute cesser de rêver à la vraie logique, et tenir les systèmes .formels pour ce qu'ils sont : des constructions, des* artefacta.

Le plan de l'ouvrage s'inspire des idées générales que je viens d'esquisser. Après un chapitre préliminaire consacré à l'examen de la définition classique de l'implication logique par la validité du conditionnel matériel, les trois chapitres suivants respectent pour l'essentiel l'ordre d'exposition adopté par Anderson et Belnap dans le volume I de la « Bible » de la logique pertinente : Entailment, The Logic of Relevance and Necessity *(Anderson et Belnap, 1975 ; 1992 pour le deuxième volume), à ceci près que l'accent est mis sur la seule pertinence plutôt que sur la pertinence conjointe à la nécessité (le système* ***E****, qui ajoute des considérations modales au souci de pertinence, ne sera évoqué qu'au passage). L'introduction progressive de fragments du système* ***R*** *me semble être la démarche la plus appropriée à la mise en valeur des motivations sous-jacentes, et à la discussion des choix successifs opérés. Le système* ***R*** *est enfin présenté au complet, l'accent étant mis, à propos de la règle Adjonction, sur les difficultés de traitement des collections de prémisses dans la perspective qui était celle d'Anderson et Belnap.* ***R*** *est ensuite muni d'une sémantique (la sémantique relationnelle de Routley et Meyer), et une preuve d'adéquation et de complétude relative à cette sémantique est présentée. La valeur et la portée de cette sémantique sont ensuite discutées, particulièrement en ce qui concerne l'interprétation qu'elle propose de la négation. Après examen et rejet de la solution « pertinentiste » proposée par Stephen Read (Read, 1988), le point de vue structural est introduit. Le chapitre VIII et dernier est enfin consacré à la question des lumières que ce point de vue peut apporter sur l'interprétation de la négation.*

Le lecteur doit être averti que bien des éléments et des résultats importants, accumulés dans l'abondante littérature sur la logique pertinente, ne sont tout simplement pas mentionnés ici (en particulier, ceux touchant aux questions de décidabilité) : les deux questions privilégiées sont celles touchant à la déduction sous hypothèses, et à la nature de la négation. Cet ouvrage n'est donc, du moins sur le plan technique, qu'une petite introduction, .même s'il prétend au statut d'essai sur le plan conceptuel. De plus, il ne concerne que le calcul propositionnel (comme l'essentiel de la *littérature pertinente, au reste)*[1]*. Ce n'est pas non plus un manuel : bien des choses sont supposées connues du lecteur, en particulier en logique classique, et d'autre part les théorèmes énoncés ne sont pas toujours accompagnés de leur démonstration. J'espère tout simplement donner envie d'en savoir plus long sur la logique pertinente, et plus encore convaincre le lecteur que la logique peut donner à penser.*

1. Dans le version publiée aux P.U.F. (2005), une annexe a été ajoutée concernant la logique pertinente quantifiée. Elle ne figure pas ici.

Quelques symboles usuels

$\rightarrow$	la conséquence : « ... implique ... », ou : « si..., alors... »
$\supset$	le conditionnel matériel, lu également : « si..., alors... »
$\wedge$, &	la conjonction extensionnelle : « ... et ... »
o	la conjonction intensionnelle, ou *fusion* : quelque chose comme : « ... avec ... »
$\vee$	la disjonction extensionnelle : « ... ou ... »
$\oplus$	la disjonction intensionnelle, ou *fission* : « … ou ... »
$\neg$, ~	la (ou les) négation(s) : « ne ... pas »
N, $\Box$	l'opérateur de nécessité : « il est nécessaire que... »
$A \vdash B$	« B est déductible de A »
$A \vDash B$	« A entraîne B » (ou « B est conséquence de A »)
<u>ssi</u>	abréviation usuelle de : « si et seulement si »

1. L'IMPLICATION LOGIQUE ET LE CONDITIONNEL

> « [Le programme est celui] d'une analyse formelle de la notion d'implication logique, diversement nommée aussi "conséquence [*entailment*]", ou "la converse de la déductibilité", telle qu'elle est exprimée par les locutions logiques comme "si..., alors ...", "implique", "entraîne", etc., celle qui correspond à des expressions logiques signalant une conclusion comme "donc", "il s'ensuit que", "d'où", "par conséquent", et autres semblables. »
>
> Anderson et Belnap, 1975, § 1. 2.

LA RELATION CLASSIQUE D'IMPLICATION LOGIQUE

L'inférence élémentaire suivante sera certainement reconnue correcte par tout un chacun :

> La Seine coule à Paris et le Danube à Vienne ;
> Donc le Danube coule à Vienne.

Et on acceptera corrélativement comme logiquement vrai l'énoncé conditionnel qui la présente sous forme implicative :

> Si la Seine coule à Paris et le Danube à Vienne,
> alors le Danube coule à Vienne.

Comment expliquer le sentiment unanime qu'une relation de conséquence logique est par deux fois exprimée ?

On fera tout d'abord remarquer que la correction de l'inférence, non plus que la vérité du conditionnel ne dépendent de « ce dont on parle », autrement dit des énoncés particuliers mis en jeu. Pour toutes les substitutions grammaticalement acceptables et uniformes d'énoncés à énoncés, les inférences obtenues demeurent correctes, de même que les conditionnels restent vrais. Ce qu'on épingle par le qualificatif « formel » accolé usuellement à « logique », et dont on prend acte en passant directement à des schémas, où des lettres représentent toutes les substitutions possibles. Schéma d'inférence valide, où le trait horizontal exprime le « donc » :

$$\frac{\text{A et B}}{\text{B}}$$

(à part quelques rares « connexivistes », peu de gens refuseraient ce schéma), ou schéma implicationnel valide (*i.e.*, dont toutes les instances possibles sont logiquement vraies) :

si A et B, alors B.

Second temps : puisque des énoncés quelconques font aussi bien l'affaire, on ne considère, à titre de *valeurs sémantiques* - *i.e.* ce qui détermine leur contribution à la validité du schéma ou à la vérité de l'énoncé conditionnel -, que les (deux, suppose-t-on) valeurs de vérité possibles, le Vrai et le Faux. On remarque alors que la validité du schéma inférentiel repose sur le fait suivant : si la conjonction A et B est vraie, alors chacun des conjoints l'est aussi, donc en particulier B, nécessairement, est vrai. Il y a conséquence ou implication logique, parce qu'il y a *transmission nécessaire de la vérité* des prémisses à la conclusion. Comment rendre compte à présent de cette nécessité ? Les combinaisons possibles des valeurs de vérités des composants donnent la solution, en termes de généralité : pour toutes les distributions de valeurs de vérité (assignations possibles des valeurs Vrai ou Faux aux lettres dans les schémas), si les prémisses ont la valeur Vrai, la conclusion aussi.

On se souvient alors qu'un schéma dont le connecteur principal est le conditionnel matériel, ou vérifonctionnel, $\supset$ (lu avec un peu de bonne volonté : « si..., alors... »), ne prend la valeur Faux que si l'antécédent est vrai, et le conséquent faux, ce qui est précisément impossible dans le cas présent. La validité manifeste, au sens intuitif, de :

si A et B, alors B,

est donc facilement expliquée en termes du concept théorique de *validité* du conditionnel matériel correspondant (désormais, « validité » veut dire : « le fait de prendre la valeur Vrai pour toutes les distributions de valeurs de vérité ») :

$(A \wedge B) \supset B.$

($\wedge$ exprimant la conjonction). Si l'on reprend à présent ce petit morceau de théorie en sens inverse, en partant des fonctions de vérité, et en particulier du conditionnel matériel, on peut penser tenir en main une explication de l'implication logique en termes vérifonctionnels : l'implication (tauto)logique, c'est la validité du conditionnel matériel.

On dira qu'en affirmant : « si A et B, alors B », j'affirmais beaucoup plus que ne le fera jamais « $(A \wedge B) \supset B$ », qui, probablement, en tant que formule d'un langage (semi)artificiel, n'affirme rien. J'affirmais quelque chose comme :

si A et B, alors, logiquement et nécessairement, B.

On accordera l'objection, mais en maintenant l'idée que ce conditionnel ordinaire est correctement rendu par le conditionnel matériel. Il suffit, pour rendre compte de la signification complète de l'affirmation, d'introduire un prédicat applicable à des schémas (ou des formules), un prédicat métalinguistique, donc, et ayant le sens de « ... est valide » :

$$\vDash$$

La validité du principe logique pris comme exemple s'exprimera donc pleinement ainsi :

$$\vDash (A \wedge B) \supset B$$

Quels sont les traits cruciaux de cette explication de la conséquence logique ? La notion d'implication (tauto)logique, en tant qu'explication de la conséquence, est d'abord vérifonctionnelle. Non pas, bien sûr, que l'implication logique soit confondue avec l'implication matérielle (ou conditionnel matériel), mais elle est fondée sur elle, via précisément la notion de validité entendue comme : « vrai pour toutes les distributions ». On est donc en présence d'une analyse essentiellement extensionnelle de la notion de conséquence logique, *i.e.* qui se dispense de faire appel aux contenus, aux sens, aux intensions exprimées. Second point : la notion d'implication logique est ici une notion métalinguistique, une relation entre des énoncés ou des schémas. On le voit clairement si on se souvient qu'elle est définie à partir de la notion de validité du conditionnel correspondant :

$$\text{A implique B si et seulement si } \vDash A \supset B.$$

Enfin, on ajoute parfois, du moins dans les variantes les plus radicales de la théorie, que, même dans des énoncés logiquement vrais, l'expression « si..., alors... » est correctement rendue par le conditionnel matériel, comme au reste dans la plupart des phrases hypothétiques[1].

Parmi toutes les explications de la notion de conséquence, cette version de l'implication logique, à côté de l'implication de Diodore, de l'implication péripatéticienne, de l'implication stricte, de l'implication intuitionniste, et de bien d'autres qu'on croisera, mérite certainement le nom de *classique.* Elle est par exemple explicitement formulée dans les *Grundlagen der Mathematik* (Hilbert et Bernays, 1934), dans le cadre des tout premiers pas vers la formalisation des inférences mathématiques[2] :

« Supposons qu'on ait montré que si A et B sont vraies toutes deux, C aussi doit être vraie. Alors il s'ensuit, selon la définition de l'implication, que l'expression :

$$A \mathbin{\&} B \supset C$$

a toujours la valeur « vraie ».

Réciproquement, si nous savons que l'expression :

$$A \mathbin{\&} B \supset C$$

en vertu des valeurs de vérité des propositions A, B, C, a la valeur « vraie », nous pouvons en conclure que, dans le cas de la vérité de A et B, C est également vraie.

> « (...) Nous reconnaissons ici l'importance qui s'attache aux expressions identiquement vraies quant au raisonnement dans le domaine de la logique des énoncés ; elles nous fournissent les schémas des suites d'inférences ; grâce à elles, les principes du raisonnement logique sont présentés en formules » (p. 120-121 de la trad. fr.)[3].

---

1. L'anglais parle de « conditionnels » là où les grammairiens français semblent parler plus volontiers de « phrases hypothétiques », voir la *Grammaire du français classique et moderne* de Wagner et Pinchon (Hachette, 1962). Le conditionnel est par ailleurs présenté comme un temps de l'indicatif. Pour éviter toute pédanterie, et au risque de m'attirer les foudres des puristes de la langue française, je suivrai quand même le modèle anglo-américain en parlant désormais des conditionnels pour les phrases en « si..., alors... ».

2 .Volume 1, § 3*b*, « Application de la théorie des fonctions de vérité à la déduction logique».

3. J'ai adapté le symbolisme utilisé par Hilbert-Bernays par souci de cohérence. Les auteurs parlent d'implication à propos de la fonction de vérité exprimée par un conditionnel matériel, et non, comme

Il faut cependant ajouter que, relativement au troisième point mentionné plus haut, les auteurs sont particulièrement circonspects : ainsi parlent-ils de « hiatus » entre l'usage de la langue, et l'interprétation de « si... , alors·· » par le conditionnel matériel :

> « Dans la langue (...), un énoncé "si A, alors B" exprime une interrelation (*Zusammenhang*), dont il suit que l'exactitude de A est une raison (*Erkenntnisgrund*) de l'exactitude de B. Le contenu de l'énoncé ne s'exprime pas simplement comme une relation entre les valeurs de vérité de A et de B » (*ibid.*, n. 1).

La reconstruction rationnelle que j'ai tentée plus haut du raisonnement qui mène à définir l'implication logique par la validité du conditionnel matériel donne à la définition vériconditionnelle de la conséquence logique un air plausible. Elle a même pour elle, peut-on dire, un parfum de sagesse des nations. Cependant elle se heurte bien vite à des difficultés devant des exemples un peu plus complexes. J'en ai déjà évoqué un, évidemment crucial, dans la *Préface*, puisque le conditionnel matériel suivant est une tautologie, *i.e.* :

$$\vDash (A \wedge \neg A) \supset B$$

Selon la définition classique de l'implication logique, on doit en conclure que $(A \wedge \neg A)$ implique logiquement B, et que l'inférence de $(A \wedge \neg A)$ à B est valide. D'un schéma d'inférence valide, toutes les instances doivent être correctes. Mais on peut hésiter sur la valeur de l'inférence suivante :

La Seine coule à Paris et la Seine ne coule pas à Paris ;
Donc Socrate est une pierre.

Autrement dit, la définition vériconditionnelle de l'implication sanctionne le principe *Ex Falso Quodlibet Sequitur* (EFQ), qui fut l'objet d'une longue discussion dans la tradition logique (on en trouve déjà une preuve en forme chez le Pseudo-Scot, mais Ockham semble le reléguer dans la catégorie des *consequentiae* non « formelles », faute sans doute de connexion nécessaire)[1]. L'objection à cette inférence serait qu'il n'y a pas le moindre lien entre les prémisses et la conclusion. Un second exemple est de nature à renforcer le doute sur la valeur de cette analyse. Le conditionnel matériel :

$$\neg(A \supset B) \supset A$$

est valide (la formule est une tautologie). On doit en conclure, au vu de la définition de l'implication logique, que :

$\neg(A \supset B)$ implique logiquement A,

et donc que toutes les inférences qui sont des instances de ce schéma sont correctes. Mais doit-on de ce fait accepter l'inférence suivante :

Il n'est pas vrai que si Dieu existe, les justes sont récompensés.
Donc Dieu existe.

Il semble au contraire qu'on puisse raisonnablement nier que l'existence de Dieu entraîne la récompense de la vertu, sans être contraint d'admettre l'existence de Dieu.

---

ici, à propos de l'implication logique. Une formule « identiquement vraie » est une tautologie du calcul propositionnel classique.

1. Cité p. 291, *in* Kneale, 1962.

À vrai dire, l'exemple est un peu plus complexe que le précédent, en raison de la présence d'un conditionnel dans la prémisse (négation de « si Dieu existe, alors les justes sont récompensés »), qu'on a rendu sans· façon par un conditionnel matériel. Sous cette traduction, peut-on objecter, la formule complète obtenue est (tautologiquement) valide, mais le désaccord avec le verdict livré par le contre-exemple provient moins de l'inadéquation de la définition de l'implication logique, que de la mauvaise traduction de la prémisse. L'inférence n'est pas concluante, bien que $\neg(A \supset B)$ implique logiquement A, mais c'est simplement que $\neg(A \supset B)$ ne rend pas vraiment compte de la prémisse de l'inférence. Il s'agirait donc moins d'une preuve d'inadéquation de la définition classique de l'implication logique, que d'une preuve de l'incapacité du conditionnel matériel à rendre compte des nombreux conditionnels qu'on trouve dans l'usage ordinaire, et souvent, comme c'est le cas ici, dans les prémisses d'une inférence.

La remarque est juste, mais est néanmoins insuffisante. *En droit,* peut-être, la question de savoir si le conditionnel matériel, défini par sa table de vérité bien connue, est capable de rendre compte des nombreux conditionnels ordinaires, est une question bien distincte de celle de la correction de la définition de l'implication logique par la validité du même conditionnel. *En fait* les deux questions se recoupent, au point de s'enchevêtrer. D'une part, parce que nous exprimons souvent (semble-t-il) des relations de conséquence par des expressions comme « si..., alors... », c'est-à-dire par des conditionnels. Mais surtout parce que la thèse selon laquelle l'implication logique est la validité du conditionnel matériel n'est réellement tenable que sur le fond d'une décision d'arrière-fond préalable, concernant la nature du langage considéré : l'extensionnalisation de toutes les connexions logiques, et donc en particulier la réduction de tous les conditionnels au conditionnel matériel.

Imaginons que nous introduisions un nouveau connecteur $\rightarrow$ pour rendre compte de la conséquence niée dans la prémisse de l'inférence (sans penser pour autant qu'il s'agit là d'une conséquence proprement logique), tout en maintenant l'idée que l'implication logique est la validité du conditionnel matériel. Nous obtenons quelque chose comme :

$$\neg(A \rightarrow B) \supset A$$

Mais comment évaluer ce conditionnel et décider de sa validité, puisque l'introduction du nouveau connecteur (clairement intensionnel) nous laisse en panne lorsqu'il s'agit de calculer les différentes valeurs de vérité que peut prendre l'antécédent ? En fait, la présence dans le langage du nouveau connecteur fait que le concept même de validité n'est plus défini, ou du moins est à redéfinir[1].

C'est pour cette (bonne) raison que toutes les discussions sur la notion d'implication, et d'implication logique en particulier, sont inextricablement liées à l'analyse du conditionnel en général. C'est le cas chez Quine, on va le voir. C'est bien sûr aussi le cas chez des adversaires « pertinentistes » : dans *Relevant Logics and their Rivals I,* par exemple, une longue partie du Chapitre I était consacrée à montrer la triple inadéquation du conditionnel matériel à capturer la notion de conséquence logique, l'implication

1. Je ne suis pas en train de dire qu'on ne peut construire un nouveau concept de validité, plus complexe, tel que cette formule soit déclarée non valide en ce nouveau sens. C'est précisément ce qu'accomplit une sémantique pour la logique pertinente, moyennant une règle d'évaluation pour $\rightarrow$. Je dis seulement que lorsqu'on définit l'implication logique par la validité du conditionnel matériel, c'est, de manière explicite ou non, relativement à un langage déjà enrégimenté extensionnellement, ou supposé tel.

nomologique, non plus que la conditionnalité proprement dite (Routley et Meyer, 1982). Au fond, il n'y a rien là que de très naturel : si une théorie adéquate de la conséquence doit être intensionnelle, c'est probablement en raison des traits d'intensionnalité que présentent les langages dans lesquels nous raisonnons.

QUINE, LOGIQUE ET GRAMMAIRE

Quine est un conservateur peu tolérant, en philosophie de la logique, et nul n'a sans doute défendu avec autant d'acharnement les valeurs de l'extensionalisme : c'est pourquoi il est utile d'examiner de manière critique son argumentaire. Je voudrais montrer, conformément à ce qui a été suggéré quelques lignes plus haut, que l'approfondissement de la réflexion logique de Quine l'a amené à mettre au jour, de plus en plus clairement, les décisions de grammaire à l'arrière-plan de la définition classique de l'implication logique. Le point de départ sera *Mathematical Logic*, le point d'arrivée *Philosophy of Logic*.

1) Dans *Mathematical Logic* (Quine, 1940), Quine énonce la définition classique de l'implication logique[1] :

> « La relation d'implication en un sens tout à fait naturel du terme, à savoir l'implication logique, peut être facilement décrite à l'aide de la notion auxiliaire de vérité logique. Un énoncé [*statement*] est logiquement vrai s'il est non seulement vrai, mais reste vrai quand tout sauf son squelette logique est varié à volonté ; en d'autres termes, s'il est vrai et ne contient de manière essentielle que des expressions logiques, les autres n'ayant que des occurrences vides. On peut dire qu'un énoncé implique logiquement un autre énoncé quand le conditionnel vérifonctionnel qui a pour antécédent le premier énoncé, et le second pour conséquent, est logiquement vrai. »[2]

L'un des exemples donnés est le suivant ; l'énoncé :

(1) si tous les hommes sont mortels, alors tous les hommes blancs sont mortels

est logiquement vrai. Pour toutes les substitutions uniformes de prédicats à « homme », « mortel », « blanc », les résultats sont des énoncés vrais. Donc l'énoncé métalinguistique ci-dessous, affirmant la relation d'implication logique entre l'antécédent et le conséquent de (1), est vrai :

(2) « tous les hommes sont mortels » implique « tous les hommes blancs sont mortels ».

Une petite question n'est cependant pas entièrement claire : comment doit-on comprendre l'expression « si..., alors... » dans l'énoncé conditionnel (1). On pourrait être tenté de le lire comme une simple variante de :

(3) tous les hommes sont mortels implique tous les hommes blancs sont mortels,

1. § 5, p. 28 de l'édition révisée de 1981.

2. Un mot figure essentiellement dans un énoncé si son remplacement par un autre mot est susceptible de transformer une vérité en une fausseté, sinon il y figure *vacuously*. Les vérités logiques sont donc les vérités où seules les particules logiques figurent de manière essentielle (Quine, 1940, p. 2).

ce que Quine, évidemment, refuse. La grammaire suffit à le condamner, car un verbe transitif comme « implique » (un prédicat binaire), exige d'être attaché à des noms, non à des énoncés, pour former un énoncé. Convaincu par ce scrupule grammatical, on pourrait néanmoins être tenté de rendre (1) par :

(4) que tous les hommes sont mortels implique que tous les hommes blancs sont mortels,

en voyant dans l'expression complexe : « que ... implique que ... », un nouveau connecteur, et en fait une simple variante grammaticale de l'expression plus familière « si..., alors... ». Mais ce n'est pas l'analyse de Quine : même dans un conditionnel logiquement vrai, l'expression « si..., alors... » doit être comprise comme un conditionnel matériel. C'est du moins ce qu'on peut conclure du fait que Quine propose, comme représentation symbolique de (1) :

(5) tous les hommes sont mortels ⊃ tous les hommes blancs sont mortels.

Autrement dit, en affirmant (1), nous affirmerions seulement que ce n'est pas le cas que tous les hommes sont mortels et qu'il y a un homme blanc non mortel. Il se trouve que l'énoncé est non seulement vrai, mais logiquement vrai, mais ce n'est pas là, à proprement parler, le contenu de notre affirmation. Nous n'affirmons la vérité logique de l'énoncé que si nous disons, justement, en parlant de lui (ou de (5)), au prix d'une *montée sémantique* :

(1) est logiquement vrai,

ou, de manière détournée mais équivalente, que son antécédent implique logiquement son conséquent. Il y a là, manifestement, une bizarrerie qu'aucune des autres remarques de ce paragraphe ne justifie.

2) Dans *Methods of Logic* (Quine, 1950)[1], l'analyse de Quine est essentiellement la même, à une légère modification près : plutôt que des énoncés logiquement vrais, on part des schémas valides, et par voie de conséquence, l'implication logique est définie en premier lieu entre schémas. Un schéma en implique un autre si, et seulement si, le conditionnel matériel avec pour antécédent le premier schéma, et pour conséquent le second, est un schéma valide. D'où la définition célèbre et rapide de l'implication logique : « L'implication est la validité du conditionnel. »

On pourrait faire remarquer que ce type de définition de la notion d'implication, à partir de celle d'énoncé logiquement vrai, a « vieilli », et que, sous l'influence notamment de Tarski, loin de définir l'implication à partir de la vérité logique, il est devenu habituel de définir d'abord la notion de conséquence, et d'en dériver, comme cas limite, la vérité logique. On dit alors qu'un ensemble quelconque Σ de prémisses implique une conclusion C, si toutes les interprétations qui satisfont les prémisses de Σ (les modèles de Σ) satisfont C. Les vérités logiques sont alors les vérités impliquées par 0 prémisse, ou, de manière équivalente, par n'importe quel ensemble de prémisses (elles sont vraies dans toutes les interprétations). Quelle différence pour la notion de conséquence ce changement de point de vue induit-il ?

1. Chap. 7, p. 49 sq., de la traduction française de l'édition de 1972.

À première vue, cela élargit considérablement la sphère d'application de la relation, du moins du côté de son domaine. Si l'on part de la vérité logique d'un conditionnel, l'antécédent ne peut contenir comme sous-formules qu'un nombre fini de formules (toutes les formules étant de longueur finie dans les langages standards), par exemple sous forme d'une conjonction. Ce qui peut se représenter :

$$(A_1 \land \ldots \land A_n) \supset C$$

D'où l'on conclut, si un tel schéma est valide, que l'ensemble *fini* de prémisses, $A_1, \ldots, A_n$, implique C. Au contraire, définir en premier lieu l'implication logique évite d'assujettir la relation à cette condition de finitude d'un ensemble de prémisses, qui revient essentiellement à une conjonction.

Conceptuellement, l'approche *à la Tarski* est évidemment préférable. Elle met en avant le fait que la logique, avant d'être un corps de vérités à côté d'autres théories portant sur des domaines particuliers, intervient comme un guide ou une norme pour les preuves dans ces théories, en disant justement « ce qui s'ensuit de quoi ». Techniquement, cependant, la différence n'est pas grande, en raison du Théorème de Compacité (qui vaut pour le Calcul propositionnel aussi bien que pour le 1^er^ ordre). Sous une de ses formes, ce Théorème dit que si un ensemble quelconque Σ, possiblement infini, implique C, *i.e.* :

$$\Sigma \vDash C,$$

alors il existe un sous-ensemble *fini* Σ_0, $\Sigma_0 \subseteq \Sigma$ tel que

$$\Sigma_0 \vDash C,$$

Il s'ensuit que si la relation de conséquence logique a lieu entre des prémisses et une conclusion, elle a toujours lieu entre un nombre fini de prémisses et la conclusion. Il n'est donc pas seulement métaphorique, dans ce cas, de parler de la conjonction (nécessairement finie) de ces prémisses. Ce n'est donc pas tellement par ce biais que la définition de Quine peut être critiquée[1].

3) Dans *Philosophy of Logic* (Quine, 1970), chap. IV, à propos des relations possibles entre énoncés d'un·langage-objet, Quine écrit :

> « Un lien familier de ce genre est *l'implication logique.* Un énoncé clos implique logiquement un autre quand, sous l'assomption que le premier est vrai, les structures des deux énoncés assurent que l'autre est vrai. La restriction cruciale ici est qu'aucun support d'une assomption supplémentaire, ou d'information, n'est invoqué quant à la vérité d'énoncés additionnels. L'implication logique repose entièrement sur la manière dont les fonctions de vérité, les quantificateurs et les variables sont combinés [*stack up*]. Elle repose entièrement sur ce que nous pouvons appeler, en un mot, la structure logique des deux énoncés. »

Pour la suite, il est plus commode de parler en termes de vérité logique, puisque la notion d'implication peut être définie à partir de celle de vérité logique : A implique B, si et seulement si A est incompatible avec non-B, si et seulement si la conjonction A et

---

1. Etchemendy semble affirmer le contraire quand il dit (Etchemendy, 1988) que la définition en termes de vérité logique du conditionnel « ne donnera jamais la pleine justification » offerte par la définition tarskienne. Pour la raison indiquée, je trouve cette affirmation équivoque, d'autant que le théorème de compacité était donné comme un corollaire du théorème de complétude de Gödel sous sa forme originelle : si $\vDash A$, alors $\vdash A$.

non-B est logiquement fausse, si et seulement si la négation de cette dernière est logiquement vraie. Cela étant, la vérité logique peut être définie en termes de structure logique des énoncés :

> « J'ai défini une vérité logique comme étant un énoncé dont la vérité est assurée par sa structure logique. Pour éviter un possible malentendu, néanmoins, il est préférable de dire plus explicitement : un énoncé est logiquement vrai si tous les énoncés qui partagent sa structure logique sont vrais. (...)
> «]'ai expliqué que ce que j'entends par la structure logique d'un énoncé, à cette étape, est sa composition en ce qui concerne les fonctions de vérité, les quantificateurs et les variables. »

C'est cette notion de structure logique constante qui permet de faire appel à la notion de substitution d'expressions à des parties composantes des énoncés, parties qu'on peut remplacer par de nouvelles, pour définir la vérité logique en termes de substitution : un énoncé logiquement vrai est un énoncé qui reste vrai pour toutes les substitutions (d'énoncés ouverts à ses énoncés simples composants, si l'on va plus loin que le calcul des énoncés). Quine observe alors qu'il y a d'autres définitions de la vérité logique (ou de la validité d'un schéma) qu'en termes de substitutions : en termes de» vrai dans toutes les interprétations sur tous les univers non vides », en termes de prouvabilité dans quelque système formel. Mais l'équivalence de la définition en termes de substitution et de la définition en termes de validité peut être démontrée, du moins pour des langages « raisonnablement riches », *i.e.* contenant l'arithmétique élémentaire :

1) Dans un sens (si un schéma est vrai pour toutes les substitutions, il est valide), le résultat est assuré par un théorème de Löwenheim-Hilbert-Bernays, disant que si un schéma A (formule du calcul du premier ordre) est satisfaisable, il est satisfaisable dans le domaine des entiers. En comprenant « un schéma est satisfaisable dans le domaine des entiers » comme « une de ses instances de substitution dans le langage de l'arithmétique est vraie », on obtient : si un schéma A est satisfaisable, il a une instance de substitution arithmétique vraie. Par contraposition, s'il est faux pour toutes ces instances, il n'a pas de modèle. Si l'on considère sa négation non-A, le résultat devient : si un schéma est vrai pour toutes les substitutions dans le langage de l'arithmétique, il est valide.

2) Pour la réciproque, par le théorème de Complétude en 1er ordre : un schéma valide est prouvable selon quelque procédure standard de preuve. Il doit s'ensuivre que toutes ses instances de substitution sont vraies.

On peut soupeser les avantages philosophiques (ontologiques en particulier) du choix de telle ou telle de ces définitions, mais le plus important, selon Quine, n'est pas le choix qu'on peut faire, mais l'équivalence de trois formulations si différentes. Cette remarquable convergence assure de la *stabilité* de la notion. C'est en ce point que Quine insiste sur le caractère fondamental de la notion de « structure logique » sous-jacente aux trois définitions :

> « Elles [les définitions du logiquement vrai] diffèrent grandement dans leur appareil conceptuel, mais elles reposent toutes sur l'identité de structure en ce qui concerne les trois constructions grammaticales particulières au langage-objet : négation, conjonction, quantification. »

Au chapitre 2 du même ouvrage, intitulé *Grammar,* Quine a déjà fait observer que la « théorie logique moderne » retient seulement dans sa grammaire logique les quatre constructions : prédication, négation, conjonction, quantification. C'est un fait, qui caractérise les langages privilégiés, ou canoniques, mais un fait rationnel ; car il s'explique et se justifie pleinement par le désir de vérifonctionnalité. C'est du moins ainsi que je comprends ces lignes du chapitre 3 :

> « Si la logique piste les conditions de vérité à travers les constructions grammaticales, et si les fonctions de vérité sont parmi ces constructions, la logique des fonctions de vérité [*truth-function logic*] est assurée. Et réciproquement, si le propos central de la logique est de pister les conditions de vérité à travers les constructions grammaticales, une grammaire artificielle construite par les logiciens est forcée d'assigner aux fonctions de vérité une place fondamentale parmi ses constructions. La grammaire que nous, logiciens, appelons tendancieusement standard est une grammaire construite avec la seule idée de faciliter la poursuite des conditions de vérité. Et c'est une très bonne idée. »

Il est important de rappeler ce point, qui souligne que Quine ne prétend pas définir, dans *Philosophy of Logic,* une notion absolue, ou, comme il dit, *transcendante,* de vérité logique. La notion, aussi diversement définie soit-elle, ne concerne qu'un langage bâti sur les fonctions de vérité et la quantification. Peut-on aller plus loin, en généralisant à des langages quelconques, sans spécifier les constructions grammaticales particulières qui permettent d'isoler la notion : énoncés ayant la même structure logique (et restant vrais pour toutes les substitutions du lexique) ? Moyennant quelques ajustements techniques, on pourrait proposer une définition comme celle-ci : « Une vérité logique est un énoncé qui ne peut être rendu faux par une substitution du lexique, même après enrichissement des ressources lexicales. »[1] Mais, fait remarquer Quine, cette notion abstraite de vérité logique ne serait pas encore transcendante, parce qu'elle dépend de ce qu'on a isolé comme construction grammaticale d'une part, comme appartenant au lexique d'autre part. Elle est donc relative à la manière dont on a construit la grammaire du langage, sous-déterminée par les données linguistiques. Même abstraite et générale, la vérité logique est « grammaire-relative ». D'où la formule célèbre de Quine : « La logique est la résultante de deux composantes : la grammaire et la vérité. »

Si la grammaire est faite de décisions théoriques en face des données disponibles, la théorie logique l'est aussi. La « grammaire canonique », comme dit *Word and Object*, celle de la prédication, de la quantification (objectuelle), et des fonctions de vérité, et la théorie logique fondée sur elle, résultent de certains choix. De quelles valeurs peuvent se recommander ces choix ? Essentiellement, semble-t-il, de la simplicité, de la transparence extensionnelle, et de l'utilité comme « schèmes pour des systèmes du monde »[2]. La logique de Quine est une logique pour la science, du moins pour une science rêvée, physicaliste et achevée. C'est bien un *organon,* mais un organon plié aux buts et aux normes de la pensée scientifique (telle, du moins, que Quine la réfléchit).

C'est pourquoi la question : la théorie logique est-elle adéquate aux inférences spontanément accomplies, et aux jugements portés sur la valeur de ces inférences ?, cette question n'a visiblement qu'un intérêt très relatif pour Quine. En cas de divergence reconnue avec l'usage, on choisira toujours l'économie et la simplicité théoriques, laissant à leur obscurité les idiomes récalcitrants. Il y a du coup un certain dogmatisme dans la façon dont Quine installe le concept véríconditionnel de validité logique. Rien

---

1. Chap. 4, p. 60.

2. Voir, par exemple, le § 47 de *Word and Object.*

qui ressemble ici à des procédures philosophiques comme l'explication carnapienne, où la question de la correspondance entre concepts formels et concepts « intuitifs » sert de règle heuristique fondamentale. Au mieux, la question de la logique étant réglée une fois pour toutes par sa mise au service du langage de la science, y a-t-il place pour des analyses subalternes, destinées à éclaircir, autant que faire se peut, les tournures de langage abandonnées en chemin. S'il y a d'autres « logiques » que la logique classique, ce n'est qu'en tant que la logique est aussi, par étymologie, étude du *logos.* L'article de *Quiddities,* « Predicate Logic », se termine par ces mots significatifs et hautains :

> « Mais le nom de logique est communément étendu aussi au-delà de ce qui peut s'ajuster à la structure grammaticale de la logique des prédicats. Nous rencontrons une logique de la nécessité et de la possibilité, une logique des questions, une logique de "tu dois" et "tu ne dois pas", des attitudes propositionnelles, et des conditionnels stricts [*strong*]. Pour autant que la grammaire de la logique des prédicats est en principe adéquate en tant que véhicule de l'austère théorie scientifique, ces entreprises ultérieures n'ont aucun rôle à jouer dans notre système théorique du monde. Encore est-il qu'elles peuvent servir de clarification pour d'autres aspects variés du langage ordinaire » (Quine, 1987).

Vu la relativité de la logique aux structures grammaticales privilégiées, soulignée par Quine, il est finalement cohérent que son analyse de l'implication passe essentiellement par une « défense et illustration » du conditionnel vérifonctionnel, de son caractère acceptable comme traduction des nombreux usages du conditionnel. Mais il me semble qu'ici le plaidoyer est particulièrement faible.

LE CONDITIONNEL

Les conditionnels irréels, ou contrefactuels, que la littérature de langue anglaise appelle aussi « conditionnels subjonctifs », sont d'emblée écartés par Quine. D'une part, ils ne sont pas concernés, et cela de manière explicite, par l'argumentation de Quine en faveur de l'analyse des conditionnels ordinaires en termes vérifonctionnels. Les seuls conditionnels visés sont les conditionnels à l'indicatif, comme :

> S'il pleut, je prends un taxi,
> Si Jean a la malaria, il a besoin de quinine,

etc. L'argument bien connu est que l'antécédent d'un conditionnel irréel étant ordinairement faux, ou supposé tel, si la table du conditionnel matériel s'appliquait, un tel conditionnel serait *ipso facto* vrai, ce qui n'est manifestement pas le cas, puisqu'on peut discuter à loisir de la valeur de vérité d'un conditionnel comme :

> Si l'ancien Premier ministre avait fait une meilleure campagne, il aurait été élu.

Ou considérons les deux exemples suivants (adaptés de Nute et Cross, 2002) :

(1) Si je faisais trois mètres de haut, j'aurais moins de 200 centimètres.
(2) Si je faisais deux mètres de haut, j'aurais plus de 150 centimètres.

Bien que les deux antécédents soient faux, et les deux conséquents vrais, le premier des deux énoncés est clairement faux, alors que le second est vrai. Le conditionnel contrefactuel n’est donc identifiable avec aucun mode de composition vérifonctionnel.

Mais - et c'est le d'« autre part » en suspens depuis trop longtemps -, cette résistance manifeste à l'analyse véríconditionnelle ne laisse pourtant pas de reste : les conditionnels intensionnels en général n'ont pas de place dans le langage austère pour la science. Ce ne sont que des expressions dramatiques, comme celles de l'idiome dispositionnel[1]. La seule question est donc de savoir dans quelle mesure le conditionnel à l'indicatif peut être analysé comme un conditionnel matériel, la thèse de Quine étant justement que « le conditionnel à l'indicatif peut toujours être construit de manière vérifonctionnelle » (Quine, 1966, § 8).

Il y a chez Quine deux arguments distincts en faveur de cette thèse, répondant à deux objections possibles.

1) Première objection : alors qu'un conditionnel à l'indicatif est bien réputé faux si l'antécédent est vrai et le conséquent faux (en accord avec la seconde ligne de la table de vérité du conditionnel matériel), il n'est pas du tout clair que tous les conditionnels ordinaires aient une valeur de vérité déterminée si l'antécédent est, ou se révèle, faux. Que je prenne ou ne prenne pas de taxi, que penser de la valeur de vérité de :

S'il pleut, je prends un taxi,

au cas où il ne pleut pas ? C'est le cas des énoncés qui expriment une réelle conditionnalité. Quine formule ce genre d'objection ainsi :

> « Une affirmation de la forme « si *p* alors *q* » est ordinairement ressentie moins comme l'affirmation d'un conditionnel que comme une affirmation conditionnelle du conséquent. Si donc, après que nous ayons avancé une telle affirmation, l'antécédent se révèle vrai, nous nous considérons comme engagés envers le conséquent et serons prêts à reconnaître notre erreur s'il se trouve faux. En revanche, s'il apparaît que l'antécédent était faux, tout se passera comme si notre affirmation conditionnelle n'avait jamais été faite » (Quine, 1950, chap. 3).

La réponse de Quine à l'objection est la suivante : il y a moins écart par rapport à l'usage ordinaire, que complétion théorique là où le verdict de l'usage n'est pas clair ou reste muet. Dans *Mathematical Logic*, Quine écrivait déjà à ce propos :

> « Il n'y a pas de conflit clair entre la table [du conditionnel matériel] et le conditionnel à l'indicatif de l'usage ordinaire. La table ne s'écarte de l'usage qu'en s'engageant sur des points où l'usage fait défaut. Cette différence est nécessaire pour toute formulation complète - toute formulation qui explique le conditionnel pour toute paire d'énoncés.
> « Ce que la table de vérité ajoute, en décidant ainsi les cas au-delà de ce que fait l'usage ordinaire, est essentiellement théorique ; aucun usage pratique supplémentaire de "si-alors" n'est par là prescrit » (Quine, 1940, § 2).

L'objection n'a du reste pas de valeur générale, dans la mesure où il y a des conditionnels à l'indicatif qui ont bien une valeur de vérité déterminée, par exemple le vrai, alors même que l'antécédent est, ou se révèle, faux. C'est le cas des conditionnels qui présentent un aspect nomologique. Regardons ce que devient la prédiction suivante :

Si les émanations de gaz carbonique augmentent, le niveau des mers montera,

au cas (miraculeux il est vrai) où les émissions de gaz carbonique dans l'atmosphère se stabiliseraient brutalement ? Il ne semble pas que la valeur de vérité de ce conditionnel soit annulée, ou suspendue, simplement parce que l'antécédent se révèle faux, puisqu'il y

1. Voir, en particulier, *The Roots of Reference*, § 3.

a tout lieu de penser que c'est justement la croyance en la vérité de ce conditionnel qui a provoqué des mesures de protection de l'environnement. Mais d'un autre côté, cet exemple ne semble pas du tout susceptible de confirmer l'adéquation de l'analyse en termes de conditionnel matériel, au contraire. Certes il a une valeur de vérité déterminée bien que l'antécédent soit (selon cette hypothèse miraculeuse) faux ; mais ce n'est pas simplement parce que l'antécédent se révèle faux que ce conditionnel peut être réputé vrai. Bien plutôt c'est la croyance en quelque connexion causale entre deux phénomènes - les émissions de gaz carbonique, l'effet de serre -, qui justifie l'affirmation de ce conditionnel. Ce point touche à la seconde objection.

2) Seconde objection : si le conditionnel à l'indicatif ordinaire exprimait une fonction de vérité, il semblerait que celui qui affirme un conditionnel, le tient pour vrai ou parce qu'il croit à la fausseté de l'antécédent, ou parce qu'il croit à la vérité du conséquent. Mais ce n'est évidemment pas le cas. Pourquoi affirmer (exemples de Quine) :

(1) Si la France est en Australie, alors l'eau de mer est douce,
(2) Si la France est en Europe, alors l'eau de mer est salée,

alors que nous sommes en mesure d'affirmer plus brièvement et plus fortement que (1) que la France n'est pas en Australie, et plus brièvement et plus fortement que (2) que l'eau de mer est salée. C'est justement quand nous ignorons les valeurs de vérité des énoncés composants que nous affirmons un conditionnel, et si nous l'affirmons, c'est donc que nous croyons en quelque lien entre ce que décrit l'antécédent et ce que décrit le conséquent. Nous pouvons nous tromper sur ce lien, mais la valeur de vérité du conditionnel en dépend, et n'est donc pas déterminée par les valeurs de vérité des énoncés composants. Donc le conditionnel à l'indicatif n'est pas une fonction de vérité. On trouve cette objection, par exemple, sous la plume de Strawson (Strawson, 1952). La réponse de Quine est plutôt elliptique :

> « En pratique, même à la lumière de la table de vérité, on ne se soucierait pas, naturellement, d'affirmer un conditionnel si on était en position d'affirmer d'entrée de jeu le conséquent, ou de nier l'antécédent - pas plus qu'on ne prend la peine d'affirmer une disjonction si l'on sait quel composant est vrai. Nous disons :
>
> Si Jean a la malaria, alors il a besoin de quinine
>
> en raison de ce que nous savons de la malaria, et parce que nous ne savons pas ce qu'il en est de l'état de santé de Jean et de son besoin de quinine. Ainsi seuls méritent d'être affirmés les conditionnels qui découlent de quelque sorte de pertinence[1] [*which follow from some manner of relevance*] entre l'antécédent et le conséquent - quelque sorte de loi, peut-être, reliant ce que les deux énoncés composants décrivent. Une telle connexion est sous-jacente aux applications utiles du conditionnel vérifonctionnel, sans participer à la signification [*meaning*] de cette notion » (Quine, 1940, § 2)[2].

Il est tout à fait surprenant de voir Quine invoquer ici une notion de signification pour asseoir sa thèse, mais nous n'en saurons pas plus, ni sur la distinction entre signification et usage (ou entre sémantique et pragmatique), ni sur les arguments qui

---

1. Sur l'histoire de l'apparition du terme « *relevance*», voir Read, 1988, p. 125

2. Il faut aller chercher paradoxalement chez Grice, 1967, la défense et illustration la plus soutenue de la thèse que le conditionnel ordinaire est bien un conditionnel matériel, quant à sa signification conventionnelle (ses conditions de vérité). Ses conditions d'assertabilité, qui restreignent son usage à certaines situations, tiennent aux maximes conversationnelles et aux contraintes qu'elles engendrent.

pourraient soutenir l'introduction de cette distinction. Mais il est clair là encore que Quine est moins préoccupé par des questions d'adéquation à l'usage, que de simplification extensionnelle, d' « enrégimentement » théorique. Son souci fondamental est de n'avoir à considérer que des modes de composition vérifonctionnels, et pour cela on confiera à la « montée sémantique » le soin d'accommoder tout ce qui déborde du cadre vérifonctionnel. Ce motif est transparent dans ce passage de *Mathematical Logic* :

> « De telles relations [comme la relation d'implication logique] sont tout à fait compatibles avec la politique consistant à supprimer les modes non vérifonctionnels de composition d'énoncés, puisqu'une relation entre énoncés n'est pas un mode de composition d'énoncés. De cette manière, la politique consistant à n'admettre que des modes de composition d'énoncés vérifonctionnels n'est pas aussi restrictive qu'il pourrait sembler à première vue ; ce qui pourrait être accompli par un conditionnel subjonctif ou un autre mode de composition non vérifonctionnel peut aussi bien l'être en parlant **au sujet** des énoncés en question, en utilisant une relation d'implication ou quelque autre forte relation entre énoncés ... » (Quine, 1940, § 5).

Ces motifs sont de stratégie théorique, pour ne pas dire de nature idéologique, et il y a quelque sophistique à présenter ces remarques comme un plaidoyer en faveur de l'*analyse* des conditionnels ordinaires par le conditionnel matériel. Mais la montée sémantique est un instrument très efficace pour purger le langage de ses traits d'intensionnalité[1] !

Or, il y a en fait des arguments très forts contre l'analyse des divers conditionnels à l'indicatif par le conditionnel matériel. Ces arguments sont d'ordre logique, en ce qu'ils opposent le comportement différent·de l'expression usuelle « si..., alors... » et du connecteur ⊃ sous certaines inférences (les exemples suivants sont tirés de Priest, 2001, et Nute et Cross, 2002). J'en donne quelques-uns en vrac, sans prétendre les analyser, puisque la logique conditionnelle n'est pas l'objet de cet ouvrage.

1) Inférences à partir de la négation d'un conditionnel : l'inférence concluant à l'existence de Dieu a déjà été prise plus haut comme exemple, je n'y reviens pas.

2) Affaiblissement de l'antécédent : il est vrai que :

$$A \supset C \vDash (A \wedge B) \supset C$$

Le conditionnel matériel n'est pas affecté par l'ajout d'une prémisse, mais ce n'est pas le cas de tous les conditionnels ordinaires :

> S'il fait beau, je sortirai ; donc s'il fait beau et je suis victime d'un accident mortel, je sortirai.

Cette inférence ressemble plutôt à une mauvaise plaisanterie. On verra au chapitre suivant que la question de la règle d'Affaiblissement est cruciale pour la logique pertinente, quand « si..., alors… » exprime une relation de déductibilité.

3) La transitivité vaut pour le conditionnel matériel :

$$A \supset B, B \supset C \vDash A \supset C$$

mais l'inférence suivante est évidemment suspecte :

1. Bien que, comme le disent justement Meyer et Sylvan, « transférer un problème au niveau du métalangage n'est pas le résoudre » (Meyer et Sylvan, 2003).

Si les autres candidats se retirent, Jean obtiendra le poste ;
si Jean obtient le poste, les autres candidats seront furieux ;
donc, si les autres candidats se retirent, ils seront furieux.

4) Toutes les variantes de contraposition sont valides en logique classique, dont :

$A \supset B \vDash \neg B \supset \neg A$

Mais l'inférence suivante laisse perplexe :

S'il pleut, les paysans sont heureux ;
donc si les paysans sont malheureux, il ne pleut pas.

Le contre-exemple suivant est particulièrement frappant :

S’il est plus de trois heures, il n'est pas beaucoup plus que trois heures ;
donc s'il est beaucoup plus que trois heures, il n'est pas plus de trois heures.

5) Conditions de négation d'un conditionnel : ·

Considérons :

(1) Si ce n'est pas Jean, c'est Paul qui l'a fait ;

Sûrs que ce n'est pas Paul le responsable, nous serons inclinés à nier ce conditionnel. Pourtant nous pouvons penser que c'est Jean qui l'a fait, et donc à quelqu'un qui affirme :

(2) Ou c'est Jean, ou c'est Paul qui l'a fait,

répondre : « Oui, l'un des deux, mais ce n'est pas Paul. » Mais si (1) était un conditionnel matériel (1) et (2) seraient équivalents et devraient appeler la même réponse (cet exemple est proposé par Stalnaker).

On pourrait bien sûr multiplier les exemples, mais la logique pertinente n'est pas essentiellement une analyse linguistique du conditionnel (bien que, selon certains, elle puisse être appliquée au conditionnel). Son objet est beaucoup plus *conceptuel* et le conditionnel ne nous intéresse ici que dans la mesure où il est susceptible d'exprimer des relations d'implication et de conséquence. Le dernier paragraphe de ce chapitre est consacré à introduire les traits fondamentaux de la logique pertinente, de ce point de vue conceptuel.

L'IMPLICATION ET LA CONSÉQUENCE PERTINENTES

Repartons de la question soulevée en page 10. Quel motif véritable a-t-on pour soutenir, comme le fait Quine, que dans l'énoncé logiquement vrai :

Si la Seine coule à Paris et le Danube à Vienne, alors la Seine coule à Paris,

l'occurrence de « si..., alors... » est celle d'un conditionnel matériel, plutôt que celle d'une expression ayant le sens de « implique » ou « entraîne logiquement » ? La seule raison invoquée par Quine est au fond un point de grammaire : « si..., alors... », comme $\supset$, est un connecteur, qui associe des énoncés pour former une nouvel énoncé composé des précédents ; dans cette composition, les énoncés composants sont utilisés, non mentionnés. Au contraire, « implique » est un verbe (ou prédicat binaire), qui exige d'être complété par des noms d'énoncés pour former un énoncé complet :

> « La Seine coule à Paris et le Danube à Vienne » implique « La Seine coule à Paris »,

noms d'énoncés formés ici avec des guillemets. Il en résulte que dans le dernier énoncé est exprimée une relation, nommément la relation d'implication logique, entre les énoncés dont il est question. Aucune telle relation n'est en jeu dans le premier, qui ne parle pas d'énoncés, mais (suppose-t-on) de Paris, de Vienne et de leurs fleuves.

D'accord pour le point de grammaire : il y a en effet une différence de construction entre l'usage d'un connecteur et celui d'un prédicat. Mais de là aux conséquences philosophiques qu'en tire Quine, il y a un grand pas qu'aucune évidence linguistique ne justifie. N'est-il pas plus naturel de soutenir que le premier énoncé veut justement dire quelque chose comme :

> Si la Seine coule à Paris et le Danube à Vienne,
> il s'ensuit [logiquement] que la Seine coule à Paris,

qui ne serait qu'une variante grammaticale du premier, où la même notion serait exprimée ? C'est précisément la *première thèse* soutenue par Anderson et Belnap, qu'à travers ces locutions variées une *même notion* est essentiellement exprimée. C'est cette notion - ou plutôt cette famille de notions -, qui doit être le centre d'intérêt de la théorie logique, qu'elle soit exprimée par un prédicat, un connecteur, ou que sais-je encore ? C'est sur la base de cet affranchissement à l'égard de la grammaire qu'a pris forme le programme originel de la logique pertinente.

On peut sommairement caractériser ce programme par trois traits fondamentaux :

1) Puisque l'implication peut être exprimée par un connecteur (et l'est effectivement dans les langues naturelles), on introduira donc formellement (*i.e.* dans le langage-objet), un nouveau connecteur $\rightarrow$, tel que $A \rightarrow B$ pourra être lu indifféremment « si A, alors B », « que A implique que B », ou même elliptiquement par un abus instruit de langage, « A implique B ». Ce connecteur a pour vocation de représenter, beaucoup plus fidèlement que $\supset$, les tours de langage présents dans nos inférences. De même qu'en logique classique les connecteurs sont caractérisables axiomatiquement, l'idée de départ était de spécifier axiomatiquement (de « définir implicitement », comme on dit parfois) les propriétés de ce nouveau connecteur d'implication. Les axiomes de Hilbert-Bernays, 1934, pour le fragment implicationnel pur du Calcul propositionnel étaient, pour mémoire, les suivants :

$A \supset (B \supset A)$	Affaiblissement
$\big(A \supset (A \supset B)\big) \supset (A \supset B)$	Contraction
$(A \supset B) \supset \big((B \supset C) \supset (A \supset C)\big)$	Transitivité

Ces axiomes, en un sens, présentent un double aspect : d'une part, en tant qu'expressions» identiquement vraies », ils» représentent en formules les principes du raisonnement logique », *i.e.* ils formalisent des inférences valides (ou supposées telles). D'autre part, en accord avec le point de vue de l'axiomatique formelle inspiré de l'exemple de la géométrie, les axiomes déterminent implicitement les propriétés des relations concernées, ici les propriétés logiques de l'implication matérielle. La recherche de « bons » axiomes pour la relation d'implication pertinente, supposée être une condition nécessaire de la validité de toute inférence, procède du même mouvement de pensée.

2) Les remarques précédentes pourraient donner à penser que la logique de la pertinence est consacrée à une seule notion, *la conséquence logique,* et il est certain qu'au départ la visée d'Anderson et Belnap était bien de capturer formellement cette notion (d'où leur préférence pour le système **E**, censé représenter la notion d'implication pertinente *et* nécessaire). Mais en fait, sous le terme générique d'implication pertinente, c'est toute une famille de notions qui sont exprimées, la conséquence proprement logique n'étant que l'une d'entre elles. Et même là, il est abusif de parler de *la notion* de conséquence logique, dans la mesure où plusieurs candidats à ce titre sont plausibles. Par exemple, le schéma $A \to (A \to A)$ est-il acceptable, ou logiquement valide (voir ci-dessous p. 41) ? Pour poser la même question autrement $(A \to A)$ est-il conséquence logique de A ? À première vue, des choix différents sont possibles et, au moins dans une perspective exploratoire, méritent d'être étudiés pour eux-mêmes[1].

Il y a donc, même en ce qui concerne la conséquence purement logique, une dose de pluralisme à l'œuvre dans la logique pertinente ; d'où l'agréable climat de tolérance dans lequel elle s'est développée.

En fait, le spectre des notions d'implication exprimées est beaucoup plus large que la seule implication logique. Dans tous les systèmes considérés, toutes les occurrences de $\to$ dans une formule sont censées exprimer une relation d'implication pertinente, dont possiblement (mais pas nécessairement) la relation de conséquence logique favorisée par tel ou tel système. Mais dans le système **R**, par exemple, seul le fait qu'une formule de la forme $A \to B$ soit *valide* (*i.e.* axiome, ou théorème d'un système, ce qu'on note « $\vdash A \to B$ ») indique que la relation de conséquence logique relative à ce système a lieu entre A et B. Dans le système modal **R□** au contraire, on peut exprimer la conséquence nécessaire par l'ajout d'un opérateur de nécessité N, et définir un nouveau connecteur d'implication nécessaire par la définition : $A \Rightarrow B =_{Df} N(A \to B)$. Naturellement, si un connecteur $\to$ comme celui de **R** n'exprime pas à lui seul l'implication logique, on est en droit de se demander quelle est au juste sa signification, et comment la représenter. On touche ici au problème d'une sémantique pour les logiques pertinentes, sur lequel on reviendra longuement.

3) Si la logique pertinente ne se caractérisait que par l'ajout d'un nouveau connecteur aux symboles logiques habituels (puisque les connecteurs extensionnels sont conservés dans les systèmes complets), elle ne serait qu'une *extension* de la logique classique, au sens où la logique modale est une extension de la logique classique par l'ajout d'un opérateur de nécessité. Mais tel n'est pas le cas : elle ne sanctionne pas comme valides toutes les inférences que la logique classique admet. Pour prendre un exemple crucial sur lequel le chapitre suivant s'attardera : la logique classique sanctionne en toute généralité le schéma d'inférence :

[1] L'ajout de ce schéma à **R** donne le système **RM**, étudié au § 29.3 d'Anderson et Belnap, 1975 ; Meyer a cependant montré qu'on pouvait dériver dans ce système $(A \to B) \vee (B \to A)$, qui donne à $\to$ une propriété de l'implication matérielle (§ 29.5). On peut en conclure que la pertinence est perdue de vue dans ce système (voir ci-dessous p. 42).

$$\frac{\text{A}}{\text{si B, alors A}}$$

qu'elle lit à l'aide du conditionnel matériel :

$$\frac{A}{(B \supset A)}$$

schéma qui correspond au « paradoxe » de l'implication matérielle, selon lequel la formule :

$$A \supset (B \supset A)$$

est valide (au sens classique de la validité). Mais imaginons qu'un mathématicien annonce, en conclusion d'un article sur la plausibilité de la conjecture de Goldbach :

> Si la conjecture de Goldbach est vraie, alors la logique du 1^er^ ordre est complète.

Il est probable qu'il susciterait le plus grand étonnement chez son auditoire. Bien sûr, il peut ajouter qu'il utilise « si..., alors... » uniquement dans le sens technique du conditionnel matériel, et que même si sa conclusion est triviale, elle est néanmoins vraie, puisque le conséquent est un énoncé mathématique vrai, et comme tel, nécessaire. À quoi on lui répondra vraisemblablement : « Soit, mais ce n'est pas le point ! Quel est le lien de conséquence entre la conjoncture de Goldbach et la complétude de la logique ? » (adapté d'Anderson et Belnap, § 3).

Cet exemple motive clairement le rejet par la logique pertinente du schéma :

$$A \rightarrow (B \rightarrow A),$$

pour des raisons, justement, de faute de pertinence : déduire de la vérité de A, ou de sa vérité nécessaire, que si B, alors A (pour n'importe quelle formule B), c'est même la faute par excellence de pertinence (on verra en revanche au chapitre 4 que la logique pertinente accepte $A \rightarrow (B \supset A)$, en tant qu'abréviation de $A \rightarrow (\neg B \vee A)$).

Un exemple encore plus frappant est celui du syllogisme disjonctif. La logique classique fournit une définition de l'implication logique telle que :

$$A \wedge (\neg A \vee B) \vDash B$$

alors que la logique pertinente rejette la validité de la formule correspondante, le Syllogisme disjonctif (SD) :

$$A \wedge (\neg A \vee B) \rightarrow B$$

pour des raisons, également, de faute de pertinence. En ce sens, la logique pertinente est une logique déviante, plutôt qu'une extension de la logique classique ; et, en fait, la situation est un peu plus complexe, dans la mesure où, bien qu'elle rejette certaines inférences, elle contient toutes les tautologies classiques comme théorèmes. Mais il n'y a rien là de surprenant : on peut à la fois admettre les formules classiquement valides, et refuser de définir l'implication logique à partir de cette notion de validité[1].

Appendice au chapitre 1
Un intermède de grammaire logique

(le lecteur pressé peut sauter cet Appendice)

La thèse fondamentale de la « Propédeutique grammaticale » (Anderson et Belnap, 1975) est la suivante : un même concept peut être représenté par toute une famille de constructions grammaticales, plus ou moins simples ou plus ou moins complexes. Appliquée à la distinction : conditionnel (« si..., alors... »), *versus* implication, la thèse nous propose de repérer et d'analyser les différentes constructions grammaticales susceptibles d'exprimer un concept générique qu'on peut appeler d' » implication-conditionnel ».

> « Thèse 1 : un concept (ou une famille de concepts) peut être isolé sans que soit spécifiée sa grammaire ; et donc on peut s'interroger sur les avantages et les inconvénients qu'il y a à l'exprimer dans différentes formes grammaticales » (Anderson et Belnap, 1975, p. 482)

QUATRE FONCTEURS FONDAMENTAUX

La grammaire logique reconnaît en général trois catégories fondamentales : celle des énoncés (ouverts ou clos), celle des termes (généreusement compris : un énoncé nominalisé comme « que la gauche a été battue aux élections », une description faisant référence au contenu d'une phrase, comme « ce que le Président a dit », sont des termes), enfin celle des *foncteurs*, entendus comme des *fonctions grammaticales*. Un foncteur transforme un ou des éléments d'une catégorie en un élément d'une catégorie. Par exemple, les prédicats sont des foncteurs qui, appliqués à un terme, donnent un énoncé. Si l'on omet les foncteurs qui peuvent s'appliquer à d'autres foncteurs (comme les adverbes qui, appliqués à un prédicat (foncteur), donnent un foncteur (prédicat composé)), et ceux qui exigent plusieurs arguments de différentes catégories, il reste quatre possibilités pour les foncteurs *élémentaires purs*, représentées dans le tableau suivant, avec les noms donnés en propre à ces foncteurs selon les cas :

Arguments	*Valeurs*	*Noms*	*Exemples*
termes	terme	Opérateur	... + ...
termes	énoncé	Prédicat	... est rouge
énoncés	énoncé	Connecteur	... et ... , Jean croit que ...
énoncés	terme	Subnecteur	« ... », que ...

1. Le connecteur extensionnel ⊃ peut être retenu à titre d'abréviation : A ⊃ B abrège ¬A ∨ B. Le point est que l' » implication » matérielle, du point de vue de la logique pertinente, n'est pas une sorte d'implication.

J'ai traduit littéralement le terme anglais « subnector » faute de mieux, en suivant le patron proposé par « connecteur » : à côté du lien créé par coordination, les subnecteurs créent un lien par subordination, du moins dans le cas clair du « that-subnector », où « que ... » est complété par une proposition subordonnée au sens des grammairiens. L'autre subnecteur est le foncteur de citation, par mise entre guillemets (voir l'analyse de la fonction de citation envisagée, puis finalement récusée par Tarski au § 1 du « Wahrheitsbegriff » (Woodger 1956)).

Des foncteurs peuvent naturellement s'appliquer à des termes déjà formés par application de foncteurs. Le prédicat (foncteur) « ... est vrai », par exemple, peut être appliqué à une expression comme « "la neige est blanche" », mais aussi à une expression comme « que la neige est blanche ». On parlera dans le premier cas de *terme d'énoncé,* dans le second de *terme de proposition.* En accord avec cette distinction, on dira que dans :

> "la neige est blanche" est vrai,

le prédicat de vérité est un prédicat d'énoncé, alors que dans :

> que la neige est blanche est vrai,

le même prédicat est un prédicat de proposition (la première construction, quoique rare sans doute hors des manuels de logique, est grammaticale ; c'est pourquoi la grammaire seule ne peut trancher entre l'hypothèse que le prédicat de vérité est un prédicat sémantique, et celle selon lequel il s'agit d'un prédicat propositionnel). De la même manière que pour les prédicats, on peut distinguer les subnecteurs d'énoncé, qui s'appliquent à un (ou des) énoncé(s) pour former un terme d'énoncé, comme le font les guillemets de citation, exemple :

> "si..., alors...",

et les subnecteurs de proposition, qui s'appliquent à un (ou des) énoncé(s) pour former un terme de proposition :

> que si..., alors...

APPLICATION À QUELQUES CONCEPTS LOGIQUES

On se demande à présent comment la grammaire logique esquissée plus haut ouvre un certain nombre de possibilités quant à l'expression d'un même concept. Prenons à titre d'exemple le concept de vérité ; les quatre sortes de foncteurs distingués plus haut font que le concept de vérité doit pouvoir être exprimé moyennant l'usage d'un opérateur, d'un prédicat, d'un connecteur, ou d'un subnecteur, qui peuvent être d'énoncé ou de proposition :

Opérateur :	que (... est vrai)	(terme ⇒ terme)
Prédicat :	... est vrai	(terme ⇒ énoncé)
Connecteur :	(que ...) est vrai	(énoncé ⇒ énoncé)
Subnecteur :	« " ... " est vrai »	(énoncé ⇒ terme)

On remarque aisément que la construction la plus simple et la plus usuelle en langage ordinaire (du moins en anglais, français, etc.), pour le concept de vérité, utilise le prédicat plutôt qu'un autre foncteur. Et donc passe par une forme de nominalisation, qu'elle prenne la forme d'une mise entre guillemets (subnecteur de citation), ou de « que » préfixé devant un énoncé (subnecteur de proposition). Il en est de même pour le concept de négation-fausseté, qui induit une nominalisation : « que ... est faux », ou» ce n'est pas le cas que ... ». On verrait aisément qu'en ce qui concerne la conjonction comme concept, le moyen d'expression le plus simple est l'usage d'un connecteur.

Qu'en est-il à présent du concept conditionnel-conséquence ? Là encore les quatre sortes de foncteurs peuvent contribuer, de manière plus ou moins simple ; à son expression. Quelques exemples (avec « que » plutôt qu'avec les guillemets) :

Opérateur :	que (si ... est vrai, alors ... est vrai)	(terme ⇒ terme)
Prédicat :	si ... est vrai, alors ... est vrai, ... entraîne ...	(terme ⇒ énoncé)
Connecteur :	si. .. , alors .. .	(énoncé ⇒ énoncé)
Subnecteur :	que ((que ...) entraîne que ...)	(énoncé ⇒ terme)

La grammaire logique prévoit, peut-on dire, des possibilités. Mais elle est applicable aux langues naturelles, pour autant que certaines de ces constructions y sont attestées : ainsi, pour le concept de conséquence, on trouve aussi bien la construction avec le prédicat binaire « entraîne » que la construction avec le connecteur « si-alors ». Évidemment, cette conception suppose (c'est la thèse 1) qu’il y a du sens à cerner et isoler un concept avant de spécifier les formes grammaticales à travers lesquelles il trouve expression. En particulier, il y a un concept « conditionnel-conséquence », qui peut être aussi bien exprimé par un connecteur que par un prédicat. Mais si nous voulons l'exprimer par un prédicat comme « entraîne », il faut nominaliser des énoncés, que ce soit par « que » ou par les guillemets de citation.

Dernière remarque sur ce point. Il y a des constructions plus simples que d'autres selon les cas et les contextes, on l'a vu : l'usage d'un prédicat est le moyen le plus simple d'exprimer le concept de vérité en langage naturel, mais, dans les langages artificiels, les logiciens préfèrent exprimer la négation-fausseté par un connecteur. Pour la conséquence, c'est la construction avec un connecteur qui est la plus simple, l'usage d'un prédicat « entraîne » obligeant à la nominalisation. Dans d'autres cas, le verdict n'est pas clair. Faut-il construire « Jean croit que la terre tourne » comme issu de l'application du connecteur « Jean croit que ... » à l'énoncé « la terre tourne », ou au contraire comme provenant de l'application du prédicat « Jean croit » au terme (obtenu par nominalisation) « que la terre tourne » ? Cette liberté d'analyse que permet la grammaire est d'importance : elle montre qu'il est possible d'analyser les énoncés de croyance sans devoir nécessairement admettre des propositions *nommées* par les expressions de la forme « que ceci ou cela ».

QUE PEUT-ON INFÉRER DE LA GRAMMAIRE ?

La réponse tient en trois mots : très peu de chose ! Il se trouve que la grammaire a ses contraintes à elle, qui font, par exemple, que la conjonction est exprimée naturellement par un connecteur inséré entre deux énoncés, alors que la négation-fausseté : « ce n'est pas le cas que ... », « il est faux que ... », exige une prénominalisation de l'énoncé nié, par le subnecteur « que ». Personne, sans doute, n'irait inférer de ce seul fait grammatical que l'usage de la locution « ce n'est pas le cas que ... » nous oblige à reconnaître des propositions, ou révèle que la négation s'applique fondamentalement aux propositions, et de manière dérivée seulement aux énoncés.

Appliquées au concept (ou au thème) conditionnel-conséquence, ces considérations amènent aux conclusions suivantes. Les formes les plus simples d'expression pour ce concept sont certes le connecteur « si.., alors... », et le prédicat binaire « entraîne ». Mais ce dernier peut contribuer, moyennant une prénominalisation par « que », à la construction d'un autre *connecteur de conséquence* : « que... entraîne que... » (il s'agit bien d'un connecteur, puisque cette expression transforme des énoncés en un énoncé). Moralité :

> « Il ne s'ensuit nullement que "la conséquence est une relation" ; parler ainsi constituerait une adhésion naïve et servile à l'idiosyncrasie de la grammaire française [*English grammar*], car la conséquence peut être représentée aussi bien par un opérateur, un connecteur, ou un subnecteur, pourvu que nous permettions suffisamment de transformations en énoncés, et de nominalisations » (Anderson et Belnap, p. 490).

Comme il a été suggéré plus haut, il n'y a aucune raison de penser que l'usage de « que » dans « que A entraîne que B », nous oblige à admettre des propositions, qui seraient les entités nommées par des expressions de la forme « que A ». Nous pouvons certainement admettre des propositions si le cœur nous en dit ; mais il ne faudrait pas croire que c'est la grammaire, ou le choix d'une tournure, qui nous les dicte. Ce point sous-tend la liberté que prennent à dessein les auteurs de lire bien souvent « $A \rightarrow B$ » : « A entraîne B », plutôt que : « que A entraîne que B », bien que cette dernière formulation soit plus correcte ; et ce, bien que les lettres « A », « B », etc., soient indiscutablement des lettres schématiques pour des énoncés ou des formules, non des lettres pour des noms d'énoncés.

Cette mise au point s'adresse principalement au lecteur qui formulerait des objections préjudicielles à la logique de la pertinence, et qui protesterait en disant que parler d'un connecteur de conséquence ou d'implication enveloppe une confusion sans espoir entre usage et mention. La distinction : connecteur/prédicat est un fait, un simple fait grammatical. Elle ne mérite pas de supporter une construction philosophique.

2. LE FRAGMENT IMPLICATIONNEL PUR DE R

« Ainsi on rejettera la validité universelle d'une formule $A \rightarrow (B \rightarrow A)$, car elle tire de A la conclusion que $B \rightarrow A$, alors que la correction de A n'a rien à voir avec la question de savoir s'il y a un lien logique entre B et A. (...) La même chose vaut pour $B \rightarrow (A \rightarrow A)$, car la validité de $A \rightarrow A$ est indépendante de la correction de B. »

Wilhelm Ackermann,
« Begründung einer strengen implication ».
(Ackermann, 1956)

Comme cette citation le suggère, l'ancêtre le plus direct de la logique pertinente est le système d'axiomes proposé par Ackermann pour l'implication « forte ». Pour saisir la substance du· fragment implicationnel pur de la logique pertinente, il convient de mettre tout de suite en lumière ce qu'entraîne le rejet par Ackermann (et, ultérieurement, par toute la tradition pertinente) de la formule jugée fautive :

$A \rightarrow (B \rightarrow A)$.

Le problème est lié à une propriété fondamentale de la déduction, qui est exprimée dans un résultat qu'on démontre par ailleurs pour certains systèmes, le Théorème de la déduction. Lorsqu'il s'agit de déduire à partir d'axiomes ou de propositions déjà acceptées, une implication de la forme : « si A, alors B », il est naturel d'ajouter l'antécédent A comme nouvelle hypothèse, de déduire B des axiomes *et* de A, et de conclure : donc les axiomes impliquent que « si A, alors B ». Ce principe exprime clairement le lien tout à fait étroit entre *implication* et *déduction* (dans l'autre sens, il s'agit tout simplement du Modus Ponens : si on a « si A, alors B », et si A est vrai, alors on peut détacher, c'est-à-dire inférer, B).

Or il semble bien que le Théorème de la déduction justifie la formule contestée[1]. Voici en effet comment on pourrait argumenter :

1) de A et de B, clairement A est déductible, ce qu'on notera pour l'instant :

$A, B \vdash A$[2] ;

1. J'espère qu'il n'y aura pas de confusion si je parle du « Théorème de· la déduction » à propos du principe général sous-jacent aux divers théorèmes-proprement dits qu'on peut obtenir pour tel et tel système, plutôt qu'à propos de ces théorèmes eux-mêmes. Pour éviter la confusion, Read utilise l'expression l' » Équivalence de la déduction », qui ne me semble cependant pas particulièrement heureuse. Je préfère donc garder la terminologie familière.

2. Il y a ici un problème majeur : faut-il écrire « A et B » (conjonction des prémisses), ou « A, B » (ensemble ou suite de prémisses) ? Cette question n'est abordée que de biais dans les premiers

2) par le Théorème de la déduction, il découle de ce fait que de A, on peut déduire que B entraîne A, noté : $A \vdash B \rightarrow A$;

3) d'où par une nouvelle application du même principe, A entraîne que B entraîne A, *i.e.* :

$$A \rightarrow (B \rightarrow A).$$

Qu'est-ce qui ne va pas ? Il est facile de le voir. En 1), la formule B n'a joué aucun rôle dans la déduction : on aurait aussi bien pu déduire A de A seule. Et en 2), le Théorème de la déduction est appliqué justement à l'hypothèse B, qui n'a pas été *réellement utilisée* dans la déduction précédente. La solution passera, non pas évidemment par l'abandon de l'idée sous-jacente au Théorème de la déduction (qui ne fait qu'exprimer le lien entre conséquence et déduction), mais par la reformulation du Théorème. Le « Théorème pertinent de la déduction » sera restreint aux cas où les hypothèses qu'on « décharge » en antécédents d'implications ont été réellement utilisées dans les déductions, ce qui revient, si l'on veut, à raffiner le concept de déduction. Il est clair qu'un tel théorème ne permettra plus de déduire la formule rejetée par Ackermann, non plus d'ailleurs que la formule $B \rightarrow (A \rightarrow A)$, où B ne joue aucun rôle dans la déduction de la formule, valide par elle-même $(A \rightarrow A)$.

On pourrait présenter d'entrée de jeu ce nouveau concept de déduction, sous la forme d'un système de Déduction naturelle (ce qui sera fait ci-dessous). Par fidélité à Anderson et Belnap, et parce que les présentations axiomatiques me semblent rendre l'identification des différents systèmes plus facile, je commencerai néanmoins par la donnée des axiomes et des règles pour **R→**, la présentation axiomatique du fragment implicationnel de **R**[1]. Je reviendrai ensuite en détail sur les motivations du choix des axiomes. Mais il faut garder en tête l'idée que les axiomes sont choisis avant tout dans la perspective d'assurer le résultat suivant : là où il y a déduction pertinente, là on peut décharger les hypothèses et obtenir *in fine* des formules qui ne commettent pas de « fautes de pertinence ».

LE SYSTÈME AXIOMATIQUE R→

Le langage-objet contient une suite infinie de variables propositionnelles, ou paramètrès (les lettres *p, q, r,* etc., seront éventuellement utilisées pour représenter l'une quelconque de ces variables), et un seul connecteur (intensionnel) →, qu'on lira « si..., alors... ». Le système **R→** est une formulation axiomatique du fragment implicationnel pur de la logique de la pertinence ; il peut être caractérisé par les axiomes suivants, dont les noms conventionnels figurent à droite :

---

chapitres d'Anderson et Belnap (1975). Elle est pourtant cruciale, puisqu'elle touche à la compréhension de ce qu'est une *collection* de prémisses. On y reviendra à propos de l'étude des formulations en Calcul des séquents.

1. **R**, bien sûr, pour *relevant implication*, l'implication pertinente.

R→ 1	$A \to A$	Identité
R→ 2	$(A \to B) \to \big((C \to A) \to (C \to B)\big)$	Transitivité (Préfixe)
R→ 3	$\big((A \to (B \to C)\big) \to \big((B \to (A \to C)\big)$	Permutation
R→ 4	$\big((A \to (A \to B)\big) \to (A \to B)$	Contraction[1]

L'unique règle d'inférence est le Modus Ponens **(M.P.)** : si $A \to B$, de A, inférer B.

Une formule A dérivée à partir des axiomes par application de la règle est dite un théorème, ce qu'on note : $\vdash A$.

À titre d'exemple, voici une dérivation (une preuve formelle) de la formule :

$A \to \big((A \to B) \to B\big)$ loi d'Assertion

Exemple 1 :

1. $(A \to B) \to (A \to B)$	**R→** 1
2. $\big((A \to B) \to (A \to B)\big) \to \big(A \to ((A \to B) \to B)\big)$	**R→** 3 (Permutation)
3. $A \to \big((A \to B) \to B\big)$	par **M.P.**

(Ultérieurement, les preuves données en exemple ne seront pas écrites *in extenso* ; on se contentera d'indiquer certaines étapes, en signalant les axiomes, théorèmes déjà démontrés, et· règles, qui les justifient.)

Ce choix d'axiomes provient de Church, 1951, « The weak theory of implication ». Une autre présentation, celle d'Anderson et Belnap, motivée, comme on le verra, par le souci d'obtenir facilement un Théorème de la déduction pertinent, est la suivante (**R→**') :

R→ 1'	$A \to A$	Identité
R→ 2'	$(A \to B) \to \big((B \to C) \to (A \to C)\big)$	Transitivité (Suffixe)
R→ 3'	$\big((A \to (B \to C)\big) \to \big((B \to (A \to C)\big)$	Permutation
R→ 4'	$\big((A \to (B \to C)\big) \to \big((A \to B) \to (A \to C)\big)$	Autodistribution

En fait, les deux formulations sont équivalentes. En présence des autres axiomes, 2 peut être remplacé par 2', et réciproquement, dans la mesure où ils sont interdéductibles (en appliquant Permutation et **M.P.**). De son côté, 4 (Contraction) peut être remplacé par 4', Autodistribution, et réciproquement ; à partir de 4', on a en effet :

1. Les logiques « de la profondeur de pertinence » [*depth relevant logics*] rejettent Contraction, aux fins d'éviter la trivialisation de théories où l'on peut exprimer l'énoncé sémantique qui mène au· paradoxe de Curry (voir Meyer, Routley et Dunn, 1979, et Priest, Routley et Norman, 1989). Comme cette question touche plus directement à la paraconsistance qu'à la pertinence proprement dite, je ne l'évoquerai plus.

1. $(A \to (A \to B)) \to ((A \to A) \to (A \to B))$ Autodistribution
2. $(A \to A) \to ((A \to (A \to B)) \to (A \to B))$ par Permutation
3. $(A \to (A \to B)) \to (A \to B)$ par **M.P.**

Dans l'autre sens, à l'aide de Contraction, on obtient Autodistribution :

1. $(A \to (B \to C)) \to (B \to (A \to C))$ Permutation
2. $(B \to (A \to C)) \to ((A \to B) \to (A \to (A \to C)))$ par Préfixe
3. $(A \to (A \to C)) \to (A \to C)$ Contraction
4. $(B \to (A \to C)) \to ((A \to B) \to (A \to C))$ Transitivité
5. $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$ Transitivité

De plus ; Permutation peut être remplacée par la loi d'Assertion, et réciproquement. On a vu plus haut qu'on obtient facilement la loi d'Assertion par Permutation ; et dans l'autre sens :

1. $B \to ((B \to C) \to C)$ Assertion
2. $(B \to ((B \to C) \to C)) \to ((((B \to C) \to C) \to (A \to C)) \to (B \to (A \to C)))$ Transitivité (suffixe)
3. $(((B \to C) \to C) \to (A \to C)) \to (B \to (A \to C))$ par M.P.
4. $(A \to (B \to C)) \to (((B \to C) \to C) \to (A \to C))$ Transitivité
5. $(A \to (B \to C)) \to (B \to (A \to C))$ Transitivité

Les différents choix possibles d'axiomes sont donc « isolés », au sens où un choix n'affecte pas les autres. On a donc la proposition (Dunn, 1986) :

PROPOSITION 1. — **R→** peut être axiomatisé en utilisant Identité, et l'un des axiomes pris dans les paires {Préfixe, Suffixe}, {Autodistribution, Contraction}, {Assertion, Permutation}, un dans chaque paire.

Exemple 2 : la formule dite *Assertion spécialisée* :

$$((A \to A) \to B) \to B$$

sera discutée par la suite (p. 36) ; en voici une preuve à partir de la loi d'Assertion :

1. $(A \to A) \to (((A \to A) \to B) \to B)$ loi d'Assertion
2. $((A \to A) \to B) \to B$ par **M.P.**

Il est généralement admis que Réflexivité et Transitivité (sous une forme ou une autre) sont caractéristiques de toute relation de conséquence[1]. La propriété de Transitivité peut être justifiée à partir du lien naturel étroit entre conséquence et déductibilité : on établit en général que B est déductible de A par une suite de pas élémentaires d'inférence qui permettent de dériver de A des formules intermédiaires, jusqu'à obtenir B ; la transitivité assure qu'on a une déduction de B à partir de A. L'admission de Permutation et Contraction est caractéristique de la logique pertinente.

FORMULATION EN DÉDUCTION NATURELLE (DNR→)

Les méthodes de déduction naturelle rendent transparentes les restrictions liées à l'exigence de pertinence. Puisque l'idée fondamentale est qu'une formule A implique *réellement* B, et donc est pertinente pour B, seulement si elle est *utilisée* dans une déduction de B à partir de A (en un sens fort de « à partir de »), les hypothèses temporairement posées sont numérotées, afin qu'on puisse garder trace de leur utilisation (par exemple, lors d'une application de la règle →Élimination, l'équivalent du Modus Ponens). De plus, une hypothèse ne peut être déchargée que si elle a été réellement utilisée dans la dérivation d'une conclusion, c'est-à-dire, si cette conclusion porte la trace de l'utilisation de l'hypothèse en question. Un système d'indices souscrits, chaque nouvelle hypothèse étant accompagnée d'un nouvel indice lors de son introduction, permet de satisfaire ces exigences.

Règles pour **DNR→**

1) Introduction d'hypothèses : au cours d'une preuve, une nouvelle hypothèse peut être introduite en tête d'une preuve subordonnée, ou sous-preuve, qu'on écrit en retrait vers la droite de la preuve déjà amorcée, avec comme indice souscrit $\{k\}$, où k est nouveau pour chaque hypothèse ;

2) $A_{\mathbf{a}}$ peut être répété à l'intérieur d'une preuve, en gardant son indice **a** ;

3) $A_{\mathbf{a}}$ peut être réitéré, *i.e.* importé d'une preuve dans une preuve subordonnée, en gardant son indice **a** ;

4) **→Élimination** : de $A_{\mathbf{a}}$ et $(A \rightarrow B)_{\mathbf{b}}$, inférer $B_{\mathbf{a} \cup \mathbf{b}}$;

5) **→Introduction** : d'une preuve de $B_{\mathbf{a}}$ sous l'hypothèse $A_{\{k\}}$, inférer $(A \rightarrow B)_{\mathbf{a}-\{k\}}$, à condition que k figure dans l'indice **a** de B. La restriction assure qu'on ne décharge A

1. Bien que Strawson ait contesté A → A, au motif que « se répéter n'est pas raisonner», et bien que Geach ait soutenu que la transitivité de l'implication n'était pas universellement valide (« Entailment », *in* Geach, 1972). Sextus Empiricus rapporte que certains (probablement les Péripatéticiens) refusent « S'il fait jour, il fait jour » au motif que l'antécédent doit être contenu dans le conséquent dans une implication vraie, et que rien ne peut se contenir soi-même (Kneale, 1962, p. 129). On trouvera quelques remarques sur les difficultés d'une logique « péripatétique » dans « Operational Semantics », par R. Sylvan et V. Plumwood, *in* Brady, 2003.

que si A a été utilisée dans la preuve de B, ce dont porte témoignage le fait que l'indice de A figure dans celui de B[1].

Il est commode de visualiser les deux dernières règles ainsi :

$$\mathbf{\to E:} \quad \begin{array}{|l} (A \to B)_{\mathbf{b}} \\ \\ A_{\mathbf{a}} \\ \hline B_{\mathbf{a}\cup\mathbf{b}} \end{array} \qquad\qquad \mathbf{\to I:} \quad \begin{array}{|l} A_{\{k\}} \\ \text{—} \\ \text{—} \\ \text{—} \\ B_{\mathbf{a}} \\ \hline (A \to B)_{\mathbf{a}-\{k\}} \end{array}$$

(à condition que $k \in \mathbf{a}$)

La dernière restriction est cruciale : c'est elle qui fait que la déduction est une *déduction pertinente,* et que la conséquence (A → B)·qui en est inférée est également pertinente, puisqu'il y a un lien entre A et B : l'hypothèse A a été réellement utilisée pour dériver B.

LE THÉORÈME DE LA DÉDUCTION PERTINENT

Explicitons plus précisément les considérations heuristiques, concernant les notions de conséquence et de déductibilité, qui conduisent au choix des règles d'inférence (en déduction naturelle), et des axiomes (pour les systèmes de style Hilbert). Et, pour ce faire, partons d'un fragment implicationnel pur du calcul propositionnel *classique,* par exemple celui proposé par Hilbert et Bernays, 1934, mais réécrit avec le connecteur de conséquence à la place de « ⊃ » :

1	$A \to (B \to A)$	Affaiblissement
2	$(A \to B) \to \big((B \to C) \to (A \to C)\big)$	Transitivité
3	$\big((A \to B) \to A\big) \to A$	Loi de Pierce

Un système qui contient l'axiome 1 (et dispose du Modus Ponens) permet de déduire, sous l'hypothèse A, la formule (B → A). Si on interprète → comme « entraîne », « a pour conséquence », et toutes locutions apparentées, cela revient à dire que l'hypothèse que A est vraie suffit à faire en sorte que n'importe quelle formule B entraîne A, ou que A peut être correctement inférée de B. Il y a deux sortes d'objection à cela.

1) Sept est bien un nombre premier ; mais personne ne dira que, pour cette raison, que la neige soit rouge *entraîne* que sept est un nombre premier (ou qu'on peut déduire ce dernier énoncé du précédent). Si B est sans pertinence relativement à A, on refusera de penser que la vérité de A entraîne le fait que A s'ensuive de B. L'axiome 1 commet une faute de pertinence, en ne respectant pas l'idée qu'une véritable implication repose sur une connexion entre l'antécédent et le conséquent. De plus, une instance de l'axiome

1. **a**, **b** ont pour valeurs possibles des ensembles d'entiers (ou de chiffres) ; dans les exemples, j'omettrai par la suite les parenthèses ensemblistes.

est $(A \rightarrow A) \rightarrow (B \rightarrow (A \rightarrow A))$, d'où par détachement $B \rightarrow (A \rightarrow A)$ est un théorème de ce système. Ici une proposition arbitraire est supposée impliquer une proposition logiquement vraie, alors même que B peut n'avoir aucun lien avec A. Il y a à nouveau ici une faute de pertinence, ce qui conduit à juger cette formule inacceptable.

2) Si B est nécessaire, comme l'est par exemple $(A \rightarrow A)$, et A ordinairement vraie (contingente), on peut déduire de A, par l'axiome 1, que B nécessaire implique A, alors qu'il est plausible de penser que les conséquences d'une proposition nécessaire sont elles-mêmes nécessaires (voir la formule K en logique modale)[1]. Il y a là, non pas peut-être une faute de pertinence, mais bien une faute de modalité, pour autant qu'une implication est une implication nécessaire.

Ces deux objections conduisent à rejeter Affaiblissement. L'observation 1) conduit au système **R→** de l'implication pertinente. Si l'on prend en compte, en outre, l'observation 2), cela conduit au système **E→**, qui vise à formaliser l'implication pertinente *et* nécessaire[2]. Le système **R→** n'a pas de préoccupations de modalité.

Positivement, quelles sont les règles plausibles pour un système de l'implication· pertinente ? Les systèmes de déduction naturelle sont un bon point de départ, on l'a dit : Anderson et Belnap utilisent à cette fin un système linéaire dû à Fitch, 1952, où les hypothèses sont introduites au fur et à mesure en tête de sous-preuves[3]. On sait déjà que l'interprétation attendue de → est que $A \rightarrow B$ veut dire indifféremment : A entraîne B, ou B est déductible de A, de sorte que la conséquence $A \rightarrow B$ est vraie (ou prouvable dans un certain contexte) si et seulement si B est *déductible* de A. Il s'ensuit que :

1) Si $(A \rightarrow B)$ est prouvable, alors l'inférence de A à B est correcte ; dans ce sens, cette exigence justifie le Modus Ponens (ou la règle →Élimination en déduction naturelle) : si on a $(A \rightarrow B)$, de A inférer B.

2) Réciproquement, il est légitime de demander que, s'il y a une déduction de B à partir de l'hypothèse A, on puisse en inférer $A \rightarrow B$. Ce qui semble justifier une règle comme →Introduction, qui dit que, si l'on a une déduction :

A
—
—
B

on peut décharger A et inférer $A \rightarrow B$.

Il y a cependant ici un problème : tout dépend de ce qu'on reconnaît comme une déduction *raisonnable* de B à partir de l'hypothèse A. Les quatre règles de déduction suivantes : Répétition, Réitération, →Élimination, →Introduction, semblent à première vue acceptables :

1) Dans une preuve sous l'hypothèse A, on doit pouvoir *répéter* A, par exemple pour inférer la conséquence $(A \rightarrow A)$ par →Introduction (→I) :

1. Il s'agit de l'axiome : $N(A \rightarrow B) \rightarrow (NA \rightarrow NB)$.
2. **E** comme « entailment », conséquence logique.
3. Prawitz (1965) présente un système de déduction naturelle de type Gentzen pertinent, en restreignant la règle →introduction aux cas où B *dépend* de A, en un sens technique : A domine B et n'a pas été déchargé.

A
A répétition
(A → A) →I

2) On doit pouvoir importer une hypothèse· figurant dans une preuve, dans une sous-preuve (et dans une sous-preuve d'une sous-preuve, etc.). L'utilisation de cette règle, conjointement avec →Élimination (→E), permet par exemple de prouver la *transitivité* de la relation de conséquence :

A → B			hyp.
	B → C		hyp.
	A → B		réitération ·
		A	hyp.
		B	→E
		B → C	réitération
		C	→E
	A → C		→I
(B → C) → (A → C)			→I

L'ennui est que ces 4 règles : Répét., Réit., →E, →I, qui définissent la notion classique de déduction, permettent évidemment de prouver la formule par excellence inadmissible :

A		
	B	
	A	réitération
	B → A	→I ?
A → (B → A)		→I ?

Mais B n'a pas été utilisée pour déduire A, et on ne souhaite pas conclure que B entraîne A (sous l'hypothèse A). D'où la restriction sur l'application de →I : on ne peut décharger une hypothèse que si elle a été utilisée· dans la déduction. L'utilisation des indices dans **DNR→** permet de formuler simplement cette restriction ;·ici, il faudrait que l'indice de B figure dans l'indice de A (ce qui ne serait pas le cas), pour qu'on puisse inférer (B → A).

Vue du point de vue de la déduction naturelle, l'exigence : si B est déductible de A, (A → B) doit être prouvable, concerne essentiellement l'usage de →Introduction. Du point de vue des systèmes avec axiomes, elle revient à demander qu'on puisse démontrer un Théorème de la Déduction. Mais ici les choses se présentent de manière différente, puisque ce sont les axiomes, et non les règles d'inférence, qui sont modifiées : bien que ce point n'ait pas été explicité jusqu'ici, le concept de déduction habituel, « Officiel », est maintenu dans le cadre de la présentation axiomatique. Or dans les formulations axiomatiques usuelles, le concept de déduction est extrêmement libéral. Relativement à un système formel **S**, une déduction de C à partir d'une liste d'hypothèses $A_1, \ldots, A_n$, est une suite de formules $B_1, \ldots, B_m$, telles que $B_m = C$ (la dernière formule est dite la « conclusion »), et chaque B_j est soit un axiome de **S**, soit une des formules A_i, soit obtenue à partir de formules antérieures par une des règles : il n'est pas exigé que toutes les formules de la suite initiale soient utilisées, les mêmes formules peuvent être répétées autant de fois qu'on veut, et dans n'importe quel ordre, etc. S'il existe une telle déduction, on écrit :

$$A_1, \ldots, A_n \vdash C$$

(C est déductible de A_1, …, A_n). Une preuve ou démonstration de C est une déduction à partir de l'ensemble vide d'hypothèses, auquel cas C est un théorème de **S**.

Classiquement, on a le Théorème de la déduction :

$$\text{Si } A_1, \ldots, A_n, A \vdash C, \text{ alors } A_1, \ldots, A_n \vdash A \rightarrow C.$$

La preuve de ce théorème montre qu'il vaut pour tout système contenant le système implicationnel **H→** (« H » pour Heyting[1]), *i.e.* pour tout système formel contenant, outre le **M.P.**, les axiomes Identité, Affaiblissement et Autodistribution ; **H→** est dit pour cette raison minimal. Inversement, on l'a vu plus haut, tout système formel pour lequel vaut un Théorème de la déduction, permet de démontrer $A \rightarrow (B \rightarrow A)$ comme théorème. Il s'ensuit que si **R→** satisfaisait un tel théorème classique de la déduction, $A \rightarrow (B \rightarrow A)$ serait un théorème de **R→**, ce qu'on ne veut pas, et ce qui n'est au reste pas le cas (voir ci-dessous, p. 38).

On pourrait voir matière à objection dans le fait que **R→** ne satisfasse pas le Théorème de la déduction. Mais on peut rétorquer que le concept de déduction a été abusivement utilisé pour caractériser des suites de formules qui ne méritent nullement ce nom (certains manuels, prudents, préfèrent justement parler de « dérivations » plutôt que de « déductions »). Ironisant sur l'appellation « Théorème de la déduction », Anderson et Belnap font remarquer à ce sujet (§ 22.2.1) :

– que parmi la suite de formules dites « hypothèses » de la dérivation, on peut trouver absolument n'importe quelles formules, vraies, fausses, valides, ou n'ayant aucun lien avec ladite « conclusion ». Elles peuvent figurer autant de fois qu'on veut, et à n'importe quelle place, dans la dérivation ; pour toutes ces raisons, le terme « prémisses » qu'on leur donne n'a rien à voir avec l'usage normal de ce terme : ce dont on a besoin pour prouver logiquement une conclusion ;

– la « conclusion » peut être elle aussi n'importe quelle formule choisie parmi les formules hypothèses, ou n'importe quelle formule figurant déjà dans la dérivation, puisque aucune répétition n'est interdite ; elle ne mérite le nom de conclusion que pour autant qu'elle est la dernière de la liste (il se trouve qu'ordinairement on est assez sage pour stopper dès qu'on a pu obtenir la conclusion recherchée !) ;

– au vu de ces remarques, le seul lien entre la définition officielle d'une déduction et une véritable déduction qui *prouve* sa conclusion, est au mieux le suivant : la définition officielle laisse ouverte la possibilité qu'une liste de formules répondant au concept officiel de déduction soit en effet une *véritable déduction* d'une conséquence à partir de prémisses.

Naturellement, cette philippique n'est pas dirigée contre le *résultat* qu'est le Théorème de la déduction pour certains systèmes, où c'est le conditionnel matériel qui est en jeu. Elle s'adresse au *nom* qu'on a donné à ce résultat. En fait, le « Théorème officiel de la déduction » ni ne concerne des déductions, ni ne permet d'en inférer qu'on a prouvé des relations de conséquences. Mais il y a un autre concept de déduction, celui de déduction pertinente, qui délivre un théorème approprié de la déduction, à l'effet que si A est réellement utilisée dans une déduction de B à partir de A_1, …, A_n, A, alors

1. Voir Heyting (1930), pour l'une des premières formalisations de la logique intuitionniste.

$(A \to B)$ est prouvable à partir de $A_1, \ldots, A_n$. On définit donc ce nouveau concept de déduction pertinente :

DÉFINITION. — Une déduction de C à partir de $A_1, \ldots, A_n$, est *pertinente relativement à* A_i, si et seulement si A_i a été utilisée pour dériver C.

L'utilisation d'une hypothèse A dans une déduction peut être marquée de la manière suivante : A est marquée de *, ainsi que toute formule provenant d'une application de **M.P.** à deux formules dont l'une au moins est marquée de *. Il est alors requis, pour qu'on ait une déduction de C *pertinente relativement à* A, que C soit aussi accompagnée de * (*the star formulation*).

Théorème pertinent de la déduction

S'il y a une déduction dans $\mathbf{R}\!\to$ de C à partir des hypothèses $A_1, \ldots, A_n$, A, qui est une déduction *pertinente* relativement à A, alors il y a une déduction (également pertinente) de $A \to C$ à partir de $A_1, \ldots, A_n$, *i.e.* :

$$\text{Si } A_1, \ldots, A_n, A^* \vdash C^*, \text{ alors } A_1, \ldots, A_n, \vdash A \to C.$$

Preuve :

Soit $B_1, \ldots, B_k$ la déduction initiale où A* figure parmi les hypothèses, et où toutes les formules qui dépendent de A* sont également étoilées. On montre que pour tout B_i, si A* a été utilisé pour l'obtenir, il y a une déduction de $A \to B_i$ à partir de $A_1, \ldots, A_n$ (on a donc aussi une déduction de $A \to C$ à partir de ces mêmes hypothèses).

Cas l :

$B_i = A^*$; auquel cas $B_i \to B_i$ est un axiome, et est donc déductible de $A_1, \ldots, A_n$.

Cas 2 :

B_i est un axiome, auquel cas A* n'est pas utilisé pour le déduire (et bien sûr B_i est déductible des hypothèses $A_1, \ldots, A_n$) ; de même si B_i est une des hypothèses $A_1, \ldots, A_n$.

Cas 3 :

B_i est obtenu par **M.P.** à partir de formules B_j, $(B_j \to B_i)$, dont l'une au moins est étoilée (si aucune n'est étoilée, B_i ne dépend pas de A, et est déductible de $A_1, \ldots, A_n$). Trois cas à nouveau sont à distinguer :

3.l : les deux formules sont étoilées, et par hypothèse de récurrence, $A \to B_j$, $A \to (B_j \to B_i)$ sont toutes deux déductibles de $A_1, \ldots, A_n$. Par Autodistribution et deux applications du Modus Ponens, on a une déduction de $A \to B_i$ à partir de $A_1, \ldots, A_n$. ·

3.2 : la prémisse mineure B_j seule est étoilée, et par hypothèse de récurrence on a une déduction de $A \to B_j$; par Transitivité, on a une déduction de $A \to B_i$ à partir de $A_1, \ldots, A_n$.

3.3 : la prémisse majeure seule est étoilée, et on a donc une déduction de $A \rightarrow (B_j \rightarrow B_i)$; par Permutation et **M.P.**, on a encore une déduction de $A \rightarrow B_i$ à partir de $A_1, \ldots, A_n$.

En fait, c'est précisément aux fins de garantir ce résultat que les axiomes ont été choisis, et cette justification éclaire la nature de **R→**, dont les axiomes pouvaient jusqu'ici sembler arbitraires. Car supposons qu'au cours d'une déduction une formule B ait été déduite (au sens fort) de A*, via des étapes intermédiaires A_i. Comment pourrait-on prouver par récurrence qu'il existe une preuve de $(A^* \rightarrow A_i^*)$, pour chaque A_i^*, donc de $(A^* \rightarrow B^*)$?

Pour amorcer la récurrence, il faut manifestement l'axiome $(A \rightarrow A)$; regardons dans la suite une formule A_j provenant de A_i et $(A_i \rightarrow A_j)$; si aucune des deux prémisses n'est étoilée, c'est qu'elles ne dépendent pas de l'hypothèse A, et donc A_j non plus. On peut donc les garder sans modification dans la nouvelle preuve qu'on s'efforce de construire. Sinon, *i.e.* pour A_j avec étoile, trois cas sont possibles :

1) la prémisse mineure est étoilée, mais pas la prémisse majeure ; on a donc $(A \rightarrow A_i)$, et $(A_i \rightarrow A_j)$. Il faut l'axiome Transitivité pour pouvoir conclure· $(A \rightarrow A_j)$;

2) la prémisse majeure est étoilée, mais pas la prémisse mineure ; on a donc par hypothèse $A \rightarrow (A_i \rightarrow A_j)$. A_i peut être conservé dans la nouvelle preuve, et on a besoin de l'axiome Permutation pour inférer $A_i \rightarrow (A \rightarrow A_j)$, et de là $(A \rightarrow A_j)$;

3) enfin les deux prémisses sont étoilées, et on a besoin de l'axiome Autodistribution pour conclure encore $(A \rightarrow A_j)$.

En bref, les axiomes Identité, Transitivité, Permutation et Autodistribution, *i.e.* les axiomes d'une des versions de **R→**, assurent justement qu'on a bien le Théorème pertinent de la déduction, tel qu'on le désire.

ÉQUIVALENCE DÉDUCTIVE DE R→ ET DNR→

PROPOSITION 2. — Les axiomes de **R→** sont prouvables selon **DNR→** ; tous les théorèmes de **R→** le sont également (→Élimination joue le rôle du Modus Ponens). Donc **DNR→** contient **R→**. Un exemple de preuve d'un axiome (Permutation) en déduction naturelle :

Exemple 3 :

1.	$A \rightarrow (B \rightarrow C)_1$	hyp.
2.	$\quad B_2$	hyp.
3.	$\quad A \rightarrow (B \rightarrow C)_1$	Réitération
4.	$\qquad A_3$	hyp.
5.	$\qquad (B \rightarrow C)_{1,3}$	→E
6.	$\qquad B_2$	Réit.
7.	$\qquad C_{1,3,2}$	→E
8.	$\quad (A \rightarrow C)_{1,2}$	→I (la restriction est respectée)
9.	$B \rightarrow (A \rightarrow C)_1$	→I (même remarque)

Une dernière application de →I permet d'obtenir Permutation.

PROPOSITION 3. — Toute preuve dans **DNR→** peut être transformée en preuve dans **R→** (la démonstration n'est qu'esquissée).

Le Théorème de la déduction pertinent montre, selon l'expression de Dunn 1986, que la règle →Introduction peut être « simulée » dans le système de style Hilbert. Ce fait peut être utilisé pour montrer qu'on peut transformer par étape une preuve en déduction naturelle en preuve dans le système **R→**. Considérons en effet la preuve subordonnée la plus à l'intérieur d'une preuve en **DNR→** (voir l'exemple plus haut) ; elle peut être considérée comme l'analogue d'une déduction avec *, où les occurrences de l'indice de l'hypothèse initiale $A_{\{k\}}$ dans les indices d'autres formules indiquent que l'hypothèse initiale a été utilisée dans leur dérivation (les autres formules proviennent de Répétition, ou de Réitération). Sur le patron de la démonstration du Théorème de la déduction, on peut donc transformer cette dérivation en une nouvelle dérivation où figurent : 1) les formules dont les indices ne contiennent pas *k* (qui donc proviennent du contexte général de la preuve) ; 2) les formules de la forme $A_{\{k\}} \rightarrow B_{\mathbf{a}}$, où **a** contient *k* ; 3) les axiomes qui permettent de justifier la présence de ces dernières ($A \rightarrow A$, Permutation, Transitivité, Autodistribution). On obtient ainsi une « quasi-preuve », *i.e.* une preuve où peuvent figurer des axiomes de **R→**, et où le nombre des hypothèses a été diminué de un. L'itération de cette démarche en remontant dans la preuve initiale permet d'obtenir finalement une suite de formules dont chacune est un axiome, ou obtenue à partir d'axiomes par **M.P.**, *i.e.* un théorème. Les deux systèmes sont donc équivalents.

LE SYSTÈME E→ DE LA CONSÉQUENCE NÉCESSAIRE

Ce système est ici présenté pour mémoire, essentiellement parce qu'Anderson et Belnap conjecturaient (probablement à tort) que **E**, dont **E→** est le fragment implicationnel, représentait fidèlement la notion de conséquence logique, pertinente *et* nécessaire [*entailment*]. Une autre voie d'accès possible· à cette notion est la méthode consistant à étendre **R** avec un opérateur de nécessité, à ajouter les axiomes appropriés, et à définir la conséquence par la nécessitation de l'implication pertinente (le système **R□**). Mais les deux auteurs ajoutent que s'il s'avérait que **R□** et **E** divergeaient, il faudrait abandonner ce dernier. Or c'est précisément ce qui s'est passé, sonnant, suivant le mot de Stephen Read, le glas de **E**[1].

On se souvient que la formule $A \rightarrow (B \rightarrow A)$ a été rejetée pour deux motifs distincts : un motif de faute de pertinence, un motif de faute de modalité : il semble qu'une vérité nécessaire ne puisse avoir pour conséquence une vérité simplement contingente. Or, bien qu'elle soit de la forme $A \rightarrow (B \rightarrow A)$, la formule :

$$A \rightarrow \big((A \rightarrow A) \rightarrow A\big),$$

est un cas particulier de la loi d'Assertion, et donc un théorème de **R→**. On peut la comprendre ainsi : si A est vraie, bien que possiblement contingente, A s'ensuit de

1. La « formule de Maksimova » $(A \rightarrow (B \rightarrow C)) \wedge (B \rightarrow (A \vee C)) \rightarrow (B \rightarrow C)$ n'est pas un théorème de **E**, alors que sa traduction dans **R□**, ($A \rightarrow B$ étant traduit par $\square(A \rightarrow B)$), est un théorème de **R□** ; la formule est par ailleurs très plausible concernant la conséquence nécessaire (voir sur ce point Anderson et Belnap, 1975, § 28.1).

$A \rightarrow A$, qui est une· vérité logique, donc nécessaire. Cette conséquence est manifestement contraire à l'idée évoquée à l'instant, selon laquelle une vérité nécessaire ne peut entrainer que des vérités nécessaires. On est donc conduit à restreindre les axiomes de **R→** pour parvenir au système **E→**, qui a pour visée de formaliser la notion de conséquence nécessaire.

AXIOMES DE **E→**. — Les axiomes de **R→**, à ceci près que la loi d'Assertion est remplacée par la loi d'Assertion *restreinte* :

$$(A \rightarrow B) \rightarrow \big(((A \rightarrow B) \rightarrow C) \rightarrow C\big).$$

La loi d'Assertion restreinte dit seulement : si une *conséquence* est vraie, alors, si elle implique une formule C, C est vraie ; elle évite de dire qu'une formule quelconque, qui peut être une variable propositionnelle, peut avoir des conséquences pour conséquences. Une autre possibilité, équivalente, est de remplacer, non pas la loi d'Assertion, mais Permutation, par Permutation *restreinte* :

$$\big(A \rightarrow ((B \rightarrow C) \rightarrow D)\big) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow D)).$$

En fait, de manière analogue à ce qui se passe pour **R→**, **E→** peut être axiomatisé en prenant Identité, et l'un des axiomes au choix des paires suivantes : {Préfixe, Suffixe}, {Contraction, Autodistribution}, {Assertion restreinte, Permutation restreinte}, un dans chaque paire.

Un cas particulier d'Assertion restreinte est :

$$(A \rightarrow A) \rightarrow \big(((A \rightarrow A) \rightarrow A) \rightarrow A\big),$$

d'où, par détachement, le théorème :

$$\big((A \rightarrow A) \rightarrow A\big) \rightarrow A$$

qui ne commet pas de faute modale, exprimant une proposition vraie : si A est conséquence d'une proposition nécessaire, alors A est vrai.

FORMULATION EN DÉDUCTION NATURELLE (**DNE→**). — Les mêmes règles que pour **DNR→**, à ceci près que la règle de Réitération est *restreinte* au cas où la formule insérée dans une sous-preuve est de la forme $(A \rightarrow B)$, *i.e.* est elle-même une conséquence. On peut vérifier, par exemple, que la restriction bloque la preuve de la loi d'Assertion non restreinte.

Au vu de ce qui a été dit plus haut, si A est la conséquence logique d'une conséquence nécessaire, A doit être nécessaire. Il semble donc qu'on puisse introduire un opérateur, N (à lire : « il est nécessaire que ... ») de la manière suivante :

$$NA = (A \rightarrow A) \rightarrow A.$$

Sous cette convention d'abréviation, le théorème : $((A \rightarrow A) \rightarrow A) \rightarrow A$ peut être réécrit :

(1) $\quad NA \rightarrow A$

(si A est nécessaire, A est vrai), qui est l'axiome caractéristique du système **T** de .logique modale. Or, avec la même convention d'abréviation, le théorème de **R→**, cas particulier de la loi d'Assertion :

$$A \to ((A \to A) \to A)$$

deviendrait :

(2) $A \to NA$.

Par ailleurs, la formule dite Assertion spécialisée :

$$((A \to A) \to B) \to B$$

est aussi un théorème de **R→** (voir p. 26, exemple 2), ce qui donne comme cas particulier :

$$((A \to A) \to A) \to A,$$

i.e. par abréviation (1) :

$$NA \to A.$$

On voit que dans **R→** les distinctions modales sont effacées, puisque A et NA y sont interdéductibles. C'est aussi pourquoi la formule $A \to ((A \to A) \to A)$ est parfois appelée *Démodalisation :* **R→** n'exprime pas l'idée de conséquence nécessaire.

Dans **E→** au contraire, on a bien le théorème $((A \to A) \to A) \to A$ (voir ci-dessus), mais *Démodalisation* n'est pas un théorème (la restriction sur Réitération en déduction naturelle, par· exemple, bloque la preuve). L'introduction de l'opérateur N est donc justifiée. On peut s'assurer que N a les propriétés attendues d'un opérateur de nécessité, du moins au sens de **S4**, en montrant que sont des théorèmes de **E→** :

1) $(A \to B) \to (NA \to NB)$ (voir la formule K en logique modale).

Preuve :

	$\vdash ((A \to A) \to A) \to ((A \to B) \to ((A \to A) \to B))$	Transitivité
	$\vdash ((A \to A) \to B)) \to ((B \to B) \to ((A \to A) \to B))$	Transitivité
d'où :	$((A \to A) \to A), (A \to B) \vdash ((B \to B) \to ((A \to A) \to B))$	**M.P.** 3 fois
d'où :	$((A \to A) \to A), (A \to B) \vdash ((A \to A) \to ((B \to B) \to B))$	Permutation restreinte et **M.P.**
d'où :	$((A \to A) \to A), (A \to B) \vdash ((B \to B) \to B)$	

Deux applications du Théorème de la déduction (et éventuellement Permutation restreinte ; mais on peut aussi réaménager l'ordre des prémisses, qui est indifférent), achèvent la preuve.

2) $(A \to B) \to N(A \to B)$ une conséquence vraie est nécessairement vraie.

Preuve : cas particulier d'Assertion restreinte.

3) $NA \to NNA$ (voir l'axiome caractéristique de **S4**).

Preuve : cas particulier de 2), étant donné que la nécessité est définie en termes de conséquence ; NA étant une conséquence, elle est elle-même nécessaire, d'où NNA.

4) $NB \to ((A \to (B \to C)) \to (A \to C))$.

Cette dernière formule mérite un rapide commentaire. Ackermann (1956) a proposé le premier calcul de l'implication nécessaire équivalent (pour le fragment implicatif pur) à **E**→ ; il donne deux règles d'inférence, →E et : de $A \to (B \to C)$ et de B, inférer $A \to C$. Mais cette deuxième règle est accompagnée de la restriction suivante : à condition que B soit une « identité logique » (une formule logiquement vraie). Le motif de la restriction est le suivant : en prenant pour B et pour C une formule non théorème A, et pour A $(A \to A)$, la règle permettrait de montrer, à partir de la supposition que A et du théorème $(A \to A) \to (A \to A)$, que $(A \to A) \to A$, *i.e.* que A est cependant nécessaire. D'où la restriction dans l'antécédent : *si* B est nécessaire (NB). La substance de la règle d'Ackermann est ici exprimée par un théorème de **E**→, qui n'a qu'une règle d'inférence.

Une autre manière de montrer que la définition de N est acceptable est d'introduire N à titre de signe primitif, muni des axiomes :

N1 $NA \to A$

N2 $(A \to B) \to N(A \to B)$

N3 $NA \to \big((A \to B) \to NB\big)$

(ajoutés à ceux de **E**→), et de montrer qu'alors NA et $(A \to A) \to A$ sont interdéductibles.

FAUTES DE PERTINENCE

L'analyse conceptuelle de l'idée d'implication pertinente (et éventuellement nécessaire) est susceptible de révéler certains critères formels selon lesquels A entraîne B de manière pertinente[1]. Il s'agit ici : 1) de proposer et de justifier des critères plausibles ; 2) de montrer que **E**→ et **R**→ satisfont ces critères. Le premier système mais non le second, est également à l'abri de fautes qu'on peut appeler « fautes de modalité ». On discutera *in fine* la question de la « pertinence » de ces critères de pertinence. Un système qui admet comme théorèmes des formules qui ne répondent pas à ces conditions est coupable de fautes de pertinence (*fallacies of relevance*). L'exemple type est évidemment celui de la formule $A \to (B \to A)$.

On a déjà longuement discuté le premier critère : pour que A entraîne B de manière pertinente, il est nécessaire qu'il existe une preuve de B qui utilise A (on n'exige pas que *toute* preuve de B utilise A). Le second procède de l'idée intuitive qu'une condition nécessaire (mais non suffisante) pour que A entraîne B est qu'un « contenu de

1. Comparer avec Hughes et Cresswell, 1968 (appendice II) : « On a suggéré (…) qu'une condition supplémentaire pour que *q* soit déductible de *p* est qu'il y ait quelque connexion de "contenu" ou de "signification" entre *p* et *q*. Mais il est extrêmement difficile, pour ne pas dire impossible, de formuler précisément ce qui est requis en plus ; et insister là-dessus semble introduire gratuitement un élément de vague dans une explication par ailleurs claire et utilisable de la déductibilité, élément qui ferait qu'il est impossible de déterminer si un système formel donné est oui ou non une logique correcte de la conséquence. »

signification » soit commun à A et à B. Cette idée rend plausible le critère formel suivant : pour que A entraîne B, A et B doivent avoir au moins une variable propositionnelle en commun (condition de pertinence *faible*). **E**→ et **R**→ satisfont ce critère :

THÉORÈME 1. — Si $A \rightarrow B$ est un théorème de **E**→, A et B ont au moins une variable en commun.

Remarque : Ce fait exclut de la classe des théorèmes des schémas comme $A \rightarrow (B \rightarrow B)$, dont une instance est $p \rightarrow (q \rightarrow q)$, où l'antécédent et le conséquent n'ont pas de variable en commun.

Preuve (Anderson et Belnap, 1975, § 5.1.2) :

On considère la matrice, où les valeurs distinguées sont indiquées par * :

→	-2	-1	+1	+2
-2	+2	+2	+2	+2
-1	-2	+1	+1	+2
*+1	-2	-1	+1	+2
*+2	-2	-2	-2	+2

Les axiomes de **E**→ prennent les valeurs distinguées pour toute assignation de valeur sur les variables, et le Modus Ponens préserve cette propriété. Mais si A et B n'ont aucune variable en commun, on peut assigner + 2 à toutes les variables de A (d'où $A = +2$), +1 à toutes les variables de B (d'où $B = +1$), ·de sorte que $A \rightarrow B$ prend la valeur non distinguée -2. Donc si A et B n'ont pas de variable en commun, $A \rightarrow B$ n'est pas prouvable[1].

Cette propriété reste vraie de **R**→ (ne pas oublier que **E**→ est strictement contenu dans **R**→) :

THÉORÈME 2. — Si $A \rightarrow B$ est un théorème de **R**→, A et B ont au moins une variable en commun.

La matrice ci-dessus satisfait $\big(A \rightarrow (B \rightarrow C)\big) \rightarrow \big(B \rightarrow (A \rightarrow C)\big)$ (Permutation non restreinte), *i.e.* cette formule prend toujours des valeurs distinguées ; donc tous les théorèmes de **R**→ prennent toujours une valeur distinguée, alors la même observation que pour **E**→ reste valable.

NOTE. — La formule $(A \rightarrow A) \rightarrow (B \rightarrow B)$ n'est donc pas un théorème, bien que $A \rightarrow A$ et $B \rightarrow B$ le soient ; on peut donner la valeur +2 à l'antécédent, la valeur +1 au conséquent, auquel cas la formule a la valeur non distinguée -2. De manière générale, il n'y a pas nécessairement implication mutuelle entre tous les théorèmes.

Le critère de la variable en commun n'est bien sûr qu'une condition nécessaire, non une condition suffisante. Dans l'instance de $A \rightarrow (B \rightarrow A)$ qu'est : $p \rightarrow (q \rightarrow p)$, l'antécédent et le conséquent ont une variable en commun, mais $A \rightarrow (B \rightarrow A)$ n'est pas pour autant un théorème de **R**→. La proposition suivante permet de s'assurer justement que cette formule n'est pas un théorème de **R**→, en montrant qu'une condition plus forte est requise pour l'appartenance à la classe des théorèmes de **R**→.

1. Sur l'usage des matrices, voir par exemple l'Introduction des traducteurs à Carnap, 1947.

THÉORÈME 3. —Si A est un théorème de **R**→, toute variable ayant une occurrence dans A figure au moins une fois comme partie antécédente, et au moins une fois comme partie conséquente de A.

Définitions *:*

1) A est partie conséquente de A ;

2) si B → C est partie conséquente de A, B est partie antécédente de A, et C partie conséquente de A ;

3) si B → C est partie antécédente de A,. C est partie antécédente de A, et B partie conséquente de A.

On considère la même matrice que plus haut. Si une variable *p* ne figure que comme partie antécédente de A, on lui donne la valeur +2, et si elle ne figure que comme partie conséquente, la valeur -2 ; on assigne aux autres variables la valeur +1. On peut montrer par récurrence que pour toute sous formule B de A *:*

1) si B ne contient pas *p*, la valeur de B est +1 ;

2) si B contient *p* et est partie antécédente de A, la valeur de B est +2.

3) si B contient *p* et est partie conséquente de A, la valeur de B est -2.

Il s'ensuit que A prend la valeur -2, et n'est donc pas prouvable.

Exemple : *q* ne figure que comme partie antécédente dans $p \rightarrow (q \rightarrow p)$; si on lui donne la valeur +2, et +1 à *p*, on obtient :

+1 → (+2 → +1), d'où +1 → -2, d'où -2.

COROLLAIRE. — Aucune variable ne peut figurer qu'une seule fois dans un théorème de **R**→.

Il découle du corollaire, par exemple, que la loi de Peirce :

$\big((A \rightarrow B) \rightarrow A\big) \rightarrow A$

n'est pas un théorème de **R**→ *:* l'instance $\big((p \rightarrow q) \rightarrow p\big) \rightarrow p$ ne contient qu'une fois la variable *q.* La faute de pertinence n'est pourtant pas évidente ici. On peut cependant faire remarquer ceci à titre de justification de ce rejet *:* supposons que si A entraîne B, alors A ; pour conclure A, il faut ajouter la prémisse : or A entraîne B. Ce qui correspond à la formule :

$(A \rightarrow B) \rightarrow \Big(\big((A \rightarrow B) \rightarrow A\big) \rightarrow A\Big)$,

qui, elle, est un théorème de **R**→ (et de **E**→), étant un cas particulier d'Assertion restreinte.

COROLLAIRE. — Aucun des deux systèmes ne contient une formule A impliquée par toute formule (une *weakest formula*). En effet, si B n'a aucune variable en commun avec A, B n'implique pas A, *i.e.* B → A n'est pas un théorème, même si A l'est (comparer avec le Calcul classique, où les tautologies sont impliquées par toute formule).

La propriété suivante ne concerne que le système **E→**. On peut considérer qu'aucune variable propositionnelle n'entraîne, au sens de la conséquence nécessaire, une formule de la forme A → B, pour tous les choix possibles de A et de B, sous peine de faute de modalité (une vérité logique n'est pas réellement déductible d'une proposition contingente, pas plus qu'on ne s'appuie sur un fait contingent pour justifier une loi logique).

PROPRIÉTÉ D'ACKERMANN : $p \to (A \to B)$ n'est pas un théorème de **E→**, où p est n'importe quelle variable propositionnelle.

On considère la matrice :

$\to$	0	1	2
0	2	2	2
1	0	2	0
*2	0	0	2

Pour cette matrice, *les* théorèmes de **E**→ prennent toujours la valeur 2 ; mais pour n'importe quelle formule $p \to (A \to B)$, en donnant à *p la* valeur 1, on obtient la valeur 0, quelles que soient les valeurs de A et de B. Une telle formule n'est donc pas prouvable.

En particulier, $A \to (A \to A)$ n'est pas un théorème de **E**→ (sinon, bien sûr, $p \to (p \to p)$ le serait). Mais cette propriété ne s'étend pas à **R**→, puisque l'instance de la loi d'Assertion :

$$p \to ((p \to p) \to p)$$

(par exemple), est évidemment un théorème de **R**→, bien que l'antécédent ne contienne qu'une variable. Cependant, contrairement à ce à quoi on pourrait s'attendre au vu de cette remarque, pas plus que $A \to (B \to B)$, où la faute de pertinence est manifeste, $A \to (A \to A)$ n'est un théorème de **R**→. Ce point n'a pas été établi ici, mais on peut néanmoins le pressentir en essayant de construire une preuve en déduction naturelle, qui échoue dans la mesure où $(A \to A)$ est dérivé sans utiliser la première occurrence de A, ce qui interdit d'appliquer **→I** (voir néanmoins p. 65-66).

En fait, l'idée de pertinence a été précisée ci-dessus en deux sens assez différents : lien de signification, expliqué en termes de variables communes (condition nécessaire), et usage effectif dans les déductions (condition à la fois nécessaire et suffisante). On peut être tenté d'en conclure que l'idée de « pertinence » est irrémédiablement confuse, et que, de ce fait, vouloir caractériser les inférences valides sur la base de critères de pertinence, c'est mettre la charrue avant les bœufs. On pourrait au contraire soutenir que c'est la validité, reconnue par ailleurs, d'une inférence, qui montre qu'il y a une relation de pertinence entre les prémisses et la conclusion. Ainsi, si la « preuve » de Lewis que de A et non-A on peut inférer n'importe quelle formule se révélait concluante (voir ci-dessous), on devrait en conclure qu'une contradiction est pertinente pour toute proposition.

Stephen Read a utilisé cet argument pour disqualifier toute tentative de caractériser la notion de conséquence logique *à partir* de l'idée de pertinence (Read, 1988, chap. 6). Il propose la démarche inverse. À partir de la notion de *compatibilité* de deux formules (telle qu'exprimée par un nouveau connecteur, **o**, appelé « fusion »), la notion de conséquence logique peut être définie :

A *entraîne* B si et seulement s'il n'est pas possible que A soit vrai *fusion* B faux.

Moyennant quoi, la relation de pertinence peut être à son tour définie, comme une relation réflexive et symétrique (mais non transitive) :

> A est pertinent pour B si ni A *fusion* non-B, ni B *fusion* non-A ne peuvent être vrais tous les deux,

autrement dit : A est pertinent pour B, si ou A entraîne B, ou B entraîne A. Read commente le point ainsi :

> « Le test pour savoir si deux propositions sont logiquement pertinentes est la question de savoir si l'une entraîne l'autre. Donc, la pertinence ne peut être épinglée *avant* qu'on établisse la validité ou la conséquence. Ce qui est important, c'est qu'après avoir épinglé les conséquences correctes par une analyse soigneuse des arguments en jeu -et non par une définition générale servant de crible aux liens de signification-, il est possible d'extraire de l'examen une notion de pertinence qui s'accorde avec l'intuition selon laquelle A est (en général) pertinent pour A et (&) B, ce dernier à son tour pertinent pour B, bien que A ne soit pas (en général) pertinent pour B » (Read, 1988, p. 133).

Je n'ai cité ici cette critique de Read que pour souligner ce qu'a de problématique la coexistence de deux critères de pertinence[1]. Je reviendrai par ailleurs au chapitre 6 sur la proposition de définir explicitement la conséquence à partir de la notion de « fusion ».

QUELQUES REMARQUES SUR LE SYSTÈME RMO→

La formule $A \rightarrow (A \rightarrow A)$, qui n'est pas un théorème de **R→**, pose un problème intéressant[2]. Du premier critère de pertinence, il découle qu'une formule de la forme $A \rightarrow (B \rightarrow C)$ ne doit être un théorème que pour autant qu'il y a une dérivation de C où les deux prémisses, A et B, sont utilisées (critère parfois dit « de Church »). Le résultat négatif mentionné à l'instant montre donc, du point de vue de **R→**, que les deux éléments de la collection (i) ci-dessous, *i.e.* les deux occurrences de A, ne sont pas confondues : l'une peut être considérée comme utilisée sans que l'autre le soit. Autrement dit, dans une tentative de dérivation comme celle-ci :

(i) A, A
(ii) A
(iii) $A \rightarrow A$
(iv) ?

on a considéré la collection des prémisses comme une suite, plutôt que comme un ensemble. Comme l'écrit Robert Meyer : « Le raisonnement semble être que dans la mesure où une suite est caractérisée par son ordre, utiliser son premier élément n'est pas utiliser son second, même si à d'autres fins nous considérons les deux éléments comme identiques » (Anderson et Belnap, 1975, § 29.3.1). La morale de cette remarque est que le système **R→** considère des occurrences distinctes de la même formule comme des

---

1. Myhill a même soutenu que la réunion de ces deux critères était incohérente (« Real Implication », *in* Norman et Sylvan, 1989).Je ne poursuivrai pas l'argument.

2. Il y a une procédure de décision pour **R→**, via sa formulation en Calcul des séquents (Kripke, 1959). La construction d'un « arbre de recherche de preuve » pour $A \rightarrow (A \rightarrow A)$ échoue en un nombre fini de pas.

hypothèses distinctes : la première occurrence de A n'étant pas utilisée, on ne peut appliquer →Introduction.

Mais on peut également faire valoir que dans une telle tentative de dérivation, les deux prémisses sont bien utilisées (A est simplement utilisée deux fois), que donc $A \to (A \to A)$ satisfait le critère de Church, et peut être acceptée comme axiome de l'implication pertinente. À l'appui de cette décision, on peut apporter les arguments suivants : il est plus « naturel » de penser que des prémisses forment un ensemble qu'une suite, d'autant que Permutation garantit que changer l'ordre des prémisses dans une suite est sans effet ; enfin : « Il semble étrange qu'un argument autrement valide doive devenir invalide parce que nous avons assumé ses prémisses trop de fois, bien que ce soit précisément la sorte de conclusion à laquelle conduit une tentative sérieuse de baser la logique sur **R→** » (Meyer, *ibid.*).

Il semble donc qu'on puisse, si on le désire, accepter la formule comme axiome, ce qui, si on l'ajoute aux quatre axiomes de **R→**, donne le système **RMO→**. Cela revient à considérer qu'une formule de la forme $A \to (B \to C)$ est un théorème si l'*ensemble* (et non plus la suite) des prémisses A, B, est utilisé dans une dérivation de C à partir de ces prémisses. L'axiome $A \to (A \to A)$ est équivalent à la règle de Mélange (« Mingle ») : de $(A \to C)$ et $(B \to C)$, inférer $A \to (B \to C)$. En effet, si la règle est prise comme primitive, via $(A \to A)$ deux fois, on obtient $A \to (A \to A)$ comme théorème. D'autre part, on peut dériver la règle avec l'axiome :

1.	$A \to C$	prémisse
2.	$B \to C$	prémisse
3.	$C \to (C \to C)$	axiome
4.	$A \to (C \to C)$	transitivité sur 1, 3
5.	$(B \to C) \to ((C \to C) \to (B \to C))$	axiome
6.	$(C \to C) \to (B \to C)$	M.P. sur 2, 5
7.	$A \to (B \to C)$	transitivité sur 4, 5

On notera que l'addition de l'axiome $A \to (A \to A)$ à **R→** préserve la propriété selon laquelle si $A \to B$ est un théorème, A et B ont au moins une variable en commun ; il ne contredit donc pas le critère faible de pertinence. Pour la matrice :

→	0	1	2
0	2	2	2
*l	0	1	2
*2	0	0	2

les axiomes de **RMO→** prennent toujours une valeur distinguée, et le Modus Ponens préserve cette propriété ; mais si A et B n'ont aucune variable en commun, on peut donner à toutes les variables de A la valeur 2, à toutes celles de B la valeur 1, ce qui donne $2 \to 1 = 0$. L'une des caractéristiques de la pertinence (avoir un « contenu » en commun) est donc sauvegardée dans le nouveau système **RMO→**. Cependant, il y a de fortes raisons de ne pas poursuivre dans cette voie. Contrairement à ce qui se passe avec **R→**, l'extension du système avec tous les connecteurs tend à se rapprocher « beaucoup trop » de la logique classique. Par exemple (pour anticiper sur la suite, voir chap. 5), le connecteur **o** a dans **RM**, comme la conjonction extensionnelle, la propriété que $A \mathbf{o} A \rightarrow A$:

1.	$A \rightarrow (A \rightarrow A)$	Axiome
2.	$A \rightarrow (\neg A \rightarrow \neg A)$	Contraposition
3.	$\neg A \rightarrow (A \rightarrow \neg A)$	Permutation ·
4.	$\neg (A \rightarrow \neg A) \rightarrow A$	Contraposition
5.	$A \mathbf{o} A \rightarrow A$	Définition de **o**

La disjonction intensionnelle $\oplus$ devient de même idempotente, comme la disjonction extensionnelle. Enfin le connecteur d'implication $\rightarrow$ acquiert la propriété du conditionnel matériel que $(A \rightarrow B) \vee (B \rightarrow A)$ est un théorème (la preuve figure dans la version publiée) : s'agit-il encore d'une implication pertinente ?

NOTE. — Du fait que $A \rightarrow (A \rightarrow A)$ n'est pas un théorème de $\mathbf{R}\boldsymbol{\rightarrow}$, on peut conclure que la converse de Contraction : $(A \rightarrow B) \rightarrow (A \rightarrow (A \rightarrow B)$, n'est pas non plus un théorème de $\mathbf{R}\boldsymbol{\rightarrow}$; si c'était. le cas, au aurait $A \rightarrow (A \rightarrow A)$ comme théorème à titre de cas particulier (par Modus Ponens).

Pour toutes ces raisons, il ne sera plus question par la suite de l'extension **RM** (avec tous les connecteurs) de ce système, **RM** occupant une position intermédiaire, et difficile à définir, entre la logique pertinente et la logique classique.

3. CONSÉQUENCE ET NÉGATION

> « La négation est une notion cruciale pour comprendre le caractère et la variété des théories de l'implication, et la notion qui sépare réellement la logique classique de la logique pertinente. »
>
> Routley et Meyer, 1982, chap. 2, § 9.

Il y a certainement de quoi s'étonner à l'idée que l'opposition entre la logique classique et la logique pertinente pivoterait sur la négation. Rien, dans les considérations précédentes sur l'implication, ne pouvait le laisser soupçonner. Et, en fait, les motifs de cette surprenante affirmation ne commenceront à se faire jour qu'au chapitre suivant, avec le rejet du Syllogisme disjonctif. En attendant, on présente ici, sans trop de souci de justification, le calcul de l'implication et de la négation (indépendamment des autres connecteurs), tel qu'Anderson et Belnap l'ont construit dans l'esprit du système de l'implication forte d'Ackermann ; puis quelques résultats techniques sont mentionnés. Comme on va le voir, les axiomes et les règles proposées donnent à la négation un « air classique ». D'abord est-ce acceptable (voir p. 88-91) ? Ensuite peut-on se fier à cette impression ? De cette nouvelle question naîtra l'idée qu'il y aurait une négation proprement pertinente, la négation dite de « De Morgan ».

LE SYSTÈME R→¬

Le langage est enrichi du connecteur de négation, ¬, ou ~ (sans qu'il y ait lieu, pour l'instant, d'y voir une différence autre que typographique), et on trouve des formules de la forme, par exemple : $\neg A$, $(A \rightarrow \neg B)$, $\neg(C \rightarrow D)$, etc. (des conséquences portant sur des négations aussi bien que des négations de conséquences)[1].

AXIOMES DE **R→¬**, (et de **E→¬**) :

1) Un choix d'axiomes caractéristiques de **R→** (ou de **E→**), par exemple Identité, Suffixe, Permutation (restreinte s'il s'agit de **E→**), Contraction.

2) Les axiomes propres à la négation :

5¬.	$(A \rightarrow \neg A) \rightarrow \neg A$	Réduction (voir remarque ci-dessous)
6¬.	$(A \rightarrow \neg B) \rightarrow (B \rightarrow \neg A)$	Contraposition intuitionniste
7¬.	$(\neg\neg A \rightarrow A)$	Double négation (non intuitionniste)

1. Chez certains auteurs, ¬ exprime la négation classique, booléenne, ~ la négation de De Morgan, cette distinction étant solidaire de l'idée qu'on a bien affaire à deux négations distinctes. Dans la mesure où je désire laisser cette question ouverte, je ne reprendrai pas systématiquement à mon compte cet usage, gardant chaque fois les choix typographiques des auteurs cités.

REMARQUE. — Le premier axiome, indispensable pour **E→¬**, est en fait redondant pour **R→¬**. La loi d'Assertion non restreinte permet de le dériver :

1.	$A \to ((A \to \neg A) \to \neg A)$	Assertion
2.	$A \to (A \to \neg(A \to \neg A))$	Contraposition, Transitivité
3.	$A \to \neg(A \to \neg A)$	Contraction
4.	$(A \to \neg A) \to \neg A$	Contraposition

Exemple 1 : $(A \to \neg\neg A)$ est un théorème de **R→¬** :

1.	$(\neg A \to \neg A) \to (A \to \neg\neg A)$	par Contraposition
2.	$(A \to \neg\neg A)$	par **M.P.**

On obtient les versions en déduction naturelle si l'on ajoute aux règles de **DNR→** (ou **DNE→** selon le cas) les trois règles suivantes pour la négation :

1) ¬Introduction (pour **DNE→**) : d'une preuve de $\neg A_{\mathbf{a}}$, sous l'hypothèse $A_{\{k\}}$, inférer $\neg A_{\mathbf{a}-\{k\}}$, à condition que l'indice *k* figure dans **a** :

$$\frac{\begin{array}{c} A_{\{k\}} \\ — \\ — \\ \neg A_{\mathbf{a}} \end{array}}{\neg A_{\mathbf{a}-\{k\}}}$$

2) Contraposition :

$$\frac{\begin{array}{ll} B_{\mathbf{a}} & \\ & A_{\{k\}} \\ & — \\ & — \\ & \neg B_{\mathbf{b}} \end{array}}{\neg A_{\mathbf{a}\cup\mathbf{b}-\{k\}}}$$

à condition que *k* figure dans **b**.

3) Double négation Élimination : de $\neg\neg A_{\mathbf{a}}$ inférer $A_{\mathbf{a}}$.

On peut vérifier aisément qu'ajouter ces règles à **DNE→** revient au même qu'ajouter les trois axiomes 5¬ - 7¬. Par exemple, si on a l'axiome 5¬ dans le contexte de la déduction naturelle, l'effet de la règle ¬Introduction est obtenu :

$$\begin{array}{ll} \quad A_{\{k\}} & \\ \quad — & \\ \quad — & \\ \quad \neg A_{\mathbf{a}} & \\ A_{\{k\}} \to \neg A_{\mathbf{a}-\{k\}} & \\ (A \to \neg A) \to \neg A & \text{Réduction} \\ \neg A_{\mathbf{a}-\{k\}} & \to\text{Élimination.} \end{array}$$

Inversement, l'axiome 5¬ peut être prouvé en utilisant la règle ¬Introduction :

$(A \rightarrow \neg A)_1$	hyp.
A_2	
$\neg A_{1,2}$	
$\neg A_1$	¬Introduction
$(A \rightarrow \neg A) \rightarrow \neg A$	→Introduction

Il en est de même, en ce qui concerne **DNR→¬**, pour les deux règles et les axiomes correspondants. On peut donc considérer que les deux versions sont équivalentes en vertu de l'équivalence déjà prouvée · pour **R→** et **NR→** ; les règles étant remplacées par les axiomes correspondants dans une preuve en déduction naturelle, les sous-preuves peuvent être progressivement transformées en preuves avec axiomes selon le patron de la démonstration donnée au chapitre précédent.

Quelques théorèmes prouvés à titre d'exemple :

Exemple 2 : $(\neg A \rightarrow A) \rightarrow A$

1	$(\neg A \rightarrow A)_1$	hyp.
2.	$\neg A_2$	hyp.
3.	$(\neg A \rightarrow A)$	Réit.
4.	$A_{1,2}$	→Élimination
5.	$\neg A_3$	hyp.
6.	$\neg A_3$	Répétition
7.	$\neg\neg A_{1,2}$	Contraposition
8.	$\neg\neg A_1$	Négation Introduction
9.	A_1	Double nég. Élimination

NOTE. — Les pas 4 à 7 peuvent être abrégés sous forme de la règle dérivée Double négation Introduction : de A, inférer ¬¬A (¬¬Introduction).

Toutes les formes de Contraposition sont prouvables, y compris la forme classique :

$$(\neg A \rightarrow B) \rightarrow (\neg B \rightarrow A)$$

Exemple 3 :

1	$(\neg A \rightarrow B)_1$	hyp.
2.	$\neg B_2$	hyp.
3.	$\neg A_3$	hyp .
4.	$(\neg A \rightarrow B)_1$	
5.	$B_{3,1}$	
6.	$\neg\ \neg B_{3,1}$	¬¬Introduction
7.	$\neg\neg A_{2,1}$	Contraposition
8.	$A_{2,1}$	Double Nég. Élimination
9.	$(\neg B \rightarrow A)_1$	→Introduction

Là encore, on peut introduire la règle dérivée : de $(\neg A \to B)$, inférer $(\neg B \to A)$.

Ces quelques exemples pour montrer la force du système, et confirmer l'impression d'avoir affaire à une négation de style « classique ».

LE THÉORÈME DE REMPLACEMENT

Ce théorème vaut également pour **R→**, et ne concerne donc pas spécialement la négation. S'il n'en a pas été fait mention, c'est uniquement pour éviter des répétitions inutiles. Il montre que les deux systèmes permettent de formuler un critère d'identité : sont intersubstituables des formules interdéductibles. On note ici « F(A) » une formule contenant une ou plusieurs occurrences de A comme sous-formule.

THÉORÈME DE REMPLACEMENT. — Si $(A \to B)$ et $(B \to A)$ sont des théorèmes (on dit aussi : si A *co-implique* B, ou si A et B sont interdéductibles), et si F(B) résulte de F(A) par remplacement de une ou plusieurs occurrences de A dans F(A) par B, alors $\big(F(A) \to F(B)\big)$ et $\big(F(B) \to F(A)\big)$ sont des théorèmes.

Exemple 4 : soit $F(q) = p \to ((p \to q) \to q)$, une instance de la loi d'assertion,
et $F(\neg\neg q) = p \to \big((p \to \neg\neg q) \to q\big)$;

puisque q et $\neg\neg q$ sont interdéductibles, on a :

$(p \to q) \to (p \to \neg\neg q)$	par Préfixe et M.P.,
$\big((p \to q) \to q\big) \to \big((p \to \neg\neg q) \to q\big)$	par Préfixe et M.P.,
$\Big(p \to \big((p \to q) \to q\big)\Big) \to \Big(p \to \big((p \to \neg\neg q) \to q\big)\Big)$	par Préfixe.

Exemple 5 : soit $F(A) = (\neg A \to B) \to (\neg B \to A)$, une forme de Contraposition,

et $F(\neg\neg A) = (\neg A \to B) \to (\neg B \to \neg\neg A)$; on a :

$(\neg B \to A) \to \big((A \to \neg\neg A) \to (\neg B \to \neg\neg A)\big)$	par Suffixe,
$(\neg B \to A) \to (\neg B \to \neg\neg A))$	par Permutation et M.P.
$(\neg A \to B) \to (\neg B \to \neg\neg A)$	par Transitivité.

Preuve du théorème, par récurrence sur la complexité des formules :

a) $F(A) = A$, auquel cas $F(B) = B$: immédiat d'après l'hypothèse.

b) On suppose que c'est vrai des formules F(A) de complexité (nombre d'occurrences de connecteurs dans la formule) inférieure à n ; pour une formule de complexité n, quatre cas sont possibles :

b1) F(A) est une formule de la forme $\neg G(A)$; par hypothèse de récurrence, on a à la fois $\big(G(A) \to G(B)\big)$ et $\big(G(B) \to G(A)\big)$ comme théorèmes ; par Contraposition, on a donc par le second : $\big(\neg G(A) \to \neg G(B)\big)$, et par le premier : $\big(\neg G(B) \to \neg G(A)\big)$ comme théorèmes, *i.e.* $\big(F(A) \to F(B)\big)$ et $\big(F(B) \to F(A)\big)$ sont des théorèmes.

b2) F(A) est une formule de la forme $\big(G \to H(A)\big)$, autrement dit A n'a d'occurrence remplacée que dans le conséquent ; par hypothèse de récurrence, on a : $H(A) \to H(B)$, et par Préfixe on a $\big(G \to H(A)\big) \to \big(G \to H(B)\big)$ comme théorème, *i.e.* $\big(F(A) \to F(B)\big)$; même raisonnement à partir de $H(B) \to H(A)$.

b3) F(A) est une formule de la forme $(G(A) \rightarrow H)$, autrement dit, A n'a d'occurrence remplacée que dans l'antécédent ; par hypothèse de récurrence, on a : $G(B) \rightarrow G(A)$, et par Transitivité (Suffixe), on a donc $(G(A) \rightarrow H) \rightarrow (G(B) \rightarrow H)$ comme théorème ; même raisonnement à partir de $G(A) \rightarrow G(B)$.

b4) F(A) est une formule de. la forme $(G(A) \rightarrow H(A))$, et F(B) est obtenu en remplaçant des occurrences de A à la fois dans l'antécédent et dans le conséquent ; en utilisant successivement Préfixe et Suffixe comme dans les cas précédents, on aboutit au résultat désiré.

En introduisant le signe $\leftrightarrow$ pour» co-implique », le théorème de remplacement peut se formuler ainsi :

si $A \leftrightarrow B$, alors $F(A) \leftrightarrow F(B)$.

R→¬ EST UNE EXTENSION CONSERVATIVE DE R→

Imaginons la situation suivante : soit **S** un système pour un langage qui ne contient que → comme connecteur, et l'unique axiome :

$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$;

soit **S'** le système obtenu à partir de **S** en ajoutant la négation, ainsi que les deux axiomes $A \rightarrow \neg\neg A$, et $\neg\neg A \rightarrow A$; on peut à présent prouver dans **S'** le théorème $A \rightarrow A$, qui, bien qu'écrit avec les seules notations de l'ancien système **S**, ne pouvait pas être prouvé avant l'introduction des nouvelles notations et des nouveaux axiomes. Autrement dit, pour prouver une formule rédigée dans le langage de **S**, il faut faire un « détour » essentiel par le nouveau système **S'**. Bien qu'étant une extension de **S**, **S'** n'en est donc pas une *extension conservative*, au sens où il y a des formules rédigées dans la notation de **S** qui ne sont prouvables qu'en passant par des axiomes de **S'**.

En général : **S'** est une extension conservative de **S** s'il s'agit d'une extension possédant la propriété suivante : soit A une formule du langage de **S** ; si A est prouvable dans **S'**, A est déjà prouvable dans **S**. Ou encore : toute formule prouvable dans **S'** sans l'être dans **S** appartient au langage de **S'**.

Il est évidemment important que l'ajout de la négation et des axiomes qui la concernent, *i.e.* le système **R→¬**, constitue une extension conservative de **R→**. C'est en effet la prétention de ce dernier système de postuler ce qui suffit à caractériser les propriétés de l'implication pertinente ; et il serait inacceptable qu'une extension destinée à traiter un tout autre sujet (la négation) vienne modifier ou enrichir ces propriétés.

PROPOSITION (Anderson et Belnap, 1975, § 14.4) : **R→¬** est une extension conservative de **R→**.

Cette proposition, non démontrée ici, est une conséquence de la propriété de la sous-formule en Calcul des séquents (voir chap. 7), et de l'équivalence d'une version appropriée de ce calcul avec **R→¬**. La propriété de la sous-formule garantit que toute formule prouvable et ne contenant que → a une preuve où ne figurent que des formules qui figurent comme sous-formules de la formule prouvée, donc où la négation n'intervient pas.

On peut aussi noter que **R→¬**, peut être axiomatisé en prenant ***f*** comme constante propositionnelle du faux, $\neg$A étant alors défini comme A → ***f*** Le seul nouvel axiome (outre les 4 axiomes pour →) :

$$((A \to f) \to f) \to A,$$

en présence de Permutation et Contraction, fait le travail des axiomes propres à la négation. En effet, Contraposition devient :

$$(A \to (B \to f)) \to (B \to (A \to f)),$$

i.e. une instance de Permutation ; Réduction devient :

$$(A \to (A \to f)) \to (A \to f),$$

i.e. une instance de Contraction ; et $\neg\neg A \to A$ devient le nouvel axiome lui-même. Toutes les propriétés de la négation sont donc obtenues dans cette nouvelle version, qui est aussi une extension conservative de **R→**.

FAUTES DE PERTINENCE

On peut se demander dans quelle mesure les propriétés de **R→**, qui jouaient le rôle de critère d'adéquation à l'égard de l'idée d'implication pertinente, peuvent être étendues au nouveau système avec négation. On a en particulier les résultats suivants :

PROPOSITION 1 : Si A → B est un théorème de **R→¬**, A et B ont au moins une variable en commun.

On peut étendre les définitions de « partie conséquente », « partie antécédente » du chapitre précédent pour accommoder la négation (cette manière de parler s'éclaircira dans le cadre du Calcul des séquents), en ajoutant aux définitions précédentes :

1) si $\neg$B est partie conséquente de A, B est partie antécédente de A ;

2 *1* si $\neg$B est partie antécédente de A, B est partie conséquente de A.

PROPOSITION 2 : Si A est un théorème de **R→¬**, alors toute variable figure au moins une fois comme partie antécédente de A, et au moins une fois comme partie conséquente de A.

Exemple : on vérifiera que dans $(p \to \neg q) \to (q \to \neg p)$, la première occurrence de *q* figure comme partie conséquente, la seconde occurrence comme partie antécédente ; de même, la première occurrence de *p* figure comme partie conséquente, la seconde comme partie antécédente.

Ce point exclut des théorèmes le schéma « paradoxal » $\neg A \to (A \to B)$, dont une instance est : $\neg p \to (p \to q)$, où *q* ne figure que comme partie conséquente.

PROPOSITION 3 : Ajouter à **R→¬** l'axiome A → (A → A) (comme on l'a ajouté à **R→** pour former **RM0→**), préserve la propriété de la PROPOSITION 1 (communauté de variables). La matrice suivante :

$\to$	0	1	2	3	$\neg$
0	3	3	3	3	3
*1	0	1	0	3	1
*2	0	0	2	3	2
*3	0	0	0	3	0

donne toujours une valeur distinguée aux théorèmes ; étant donné une formule ($A \to B$) sans variable commune, donner 1 aux variables de A, 2 aux variables de B, fait qu'on obtient $1 \to 2 = 0$.

PREMIÈRES QUESTIONS SUR LA NÉGATION

Dans Anderson et Belnap, 1975, on ne trouvait pas de justification véritable des axiomes pour la Négation, Contraposition et Double Négation Élimination. Dans un article intitulé « What is Relevant Implication? », Alasdair Urquhart a soutenu que ces règles pour la négation sont incorrectes au regard de la logique pertinente. Bien que les arguments anticipent le contenu du chapitre suivant, il peut être utile d'en donner déjà une idée. Voici la thèse d'Urquhart :

> « Il est clair que dans **R** lui-même, la négation, quoique non complètement classique (puisque $A \,\&\, \neg A \to B$ fait défaut), était pensée comme un *genre* de négation classique. Par exemple, la loi du Tiers exclu et de la double négation étaient postulées pour elle.
> « Tout cela, je crois, est une erreur.
> « (...) Ce n'est pas seulement que la logique pertinente peut être intuitionniste ; elle *doit* être intuitionniste » (*in* Norman et Sylvan, 1989).

La thèse soutenue par Urquhart était que les lois du Tiers Exclu et de Double Négation reposent sur une faute de pertinence cachée. Dans le détail, l'argument n'est pas absolument clair, mais repose en gros sur l'idée que la règle de déduction naturelle pour ∨Élimination, telle qu'adoptée par Anderson et Belnap, est incorrecte. Il faut la reformuler sous forme de deux dérivations en parallèle d'une même conclusion sous chacun des disjoints (ce qui, on le verra, n'est pas exactement le cas dans la formulation d'Anderson et Belnap, qui utilise les lois de De Morgan) :

1	$(A \vee B)_{\mathbf{a}}$		
2		$A_{\mathbf{a}}$	$B_{\mathbf{a}}$
		—	—
		—	—
n		$C_{\mathbf{a}\cup\mathbf{b}}$	$C_{\mathbf{a}\cup\mathbf{b}}$
n+1	$C_{\mathbf{a}\cup\mathbf{b}}$		

Cette reformulation a pour elle que la loi de distributivité peut être prouvée, au lieu d'être simplement postulée (voir chap. 5). Cela posé, l'argument de Urquhart se développe ainsi. Définissons la négation comme on le fait en logique minimale, à l'aide de la constante du faux ***f*** *:*

$$\neg A =_{\mathrm{Def}} A \to \boldsymbol{f}$$

À quoi pourrait ressembler une dérivation du Tiers Exclu ? Selon lè format des dérivations en parallèle pour la disjonction, on aurait :

1		A	$\boldsymbol{f}$	$\rightarrow$Introduction
2			$A \rightarrow \boldsymbol{f}$	$\vee$Introduction et abréviation
3		$A \vee \neg A$	$A \vee \neg A$	$\vee$Élimination
4	$A \vee \neg A$			

Il est clair que cette preuve, admissible en logique classique, ne peut être acceptée en logique pertinente : le pas 2 est incorrect, puisque $\boldsymbol{f}$ ne dépend pas de A, ou que A n'a pas été réellement utilisé dans la dérivation de $\boldsymbol{f}$. Là est la faute grossière de pertinence, sur laquelle repose la « soi-disant preuve » du Tiers exclu[1].

Donc, la négation en logique pertinente ne peut être le genre de négation classique qu'Anderson et Belnap ont incorporée à **R**. L'analyse de Urquhart appelle d'ores et déjà deux remarques distinctes :

1) L'argument est-il de nature à justifier l'idée que le concept d'implication pertinente est « intuitionniste plutôt que classique » ? En fait, il a une allure quelque peu circulaire. On suppose qu'on doit partir du concept *minimal* de négation ($\neg A$ défini par $A \rightarrow \boldsymbol{f}$) et l'argument tend à montrer qu'une preuve du Tiers exclu à partir de ce concept contient nécessairement une faute de pertinence. On n'a nullement montré qu'un concept de négation qui pourrait justifier le Tiers exclu, est en lui-même incohérent avec l'idée d'implication pertinente - à supposer que cette idée soit claire.

2) Urquhart affirme que la négation de **R** est un *genre* de négation classique, sans autre précision. Il admet cependant qu'elle n'est pas complètement classique, vu l'absence de *Ex Falso Quodlibet.* Après quoi il conclut que la logique pertinente est, et doit être, intuitionniste. Mais le problème se repose : la logique intuitionniste admet la validité de $A \;\&\; \neg A \rightarrow B$; le système de logique pertinente que propose Urquhart ne l'admet évidemment pas. Donc il n'est pas plus *complètement* intuitionniste que **R** n'est complètement classique ? On verra plus loin quelles conclusions négatives, cette fois fondées sur des considérations sémantiques et nettement plus convaincantes, Urquhart a tiré de cet état de choses. Ces questions difficiles seront développées plus avant dans les chapitres 6 et 8.

1. Urquhart affirme qu'une analyse exactement semblable montrerait qu'il en est de même pour la preuve de $\neg\neg A \rightarrow A$, mais ne la donne pas. On peut du moins imaginer un cas particulier du Syllogisme disjonctif qui mènerait de $\neg\neg A$ à A via le Tiers exclu : $\neg\neg A, A \vee \neg A \vdash A$.

4. LA CONSÉQUENCE ENTRE FONCTIONS DE VÉRITÉ

« Nous avions été désespérés qu'on ne puisse prouver dans le système d'Ackermann le syllogisme disjonctif, de A ou B *et* non-A, inférer B, à titre d'implication forte ou rigoureuse. Alors un jour, avec toute l'hésitation nécessaire, j'avançai l'idée que c'était peut-être parce que l'argument lui-même était fautif. La réponse d'Alan fut du genre : Absurde !, ct la chose fut enterrée. Mais le jour suivant, Alan revint et : "Bon, peut-être ... " »

Belnap, « Dedicatory note on Alan Anderson » (1974), *in* Norman et Sylvan, 1989.

Le langage est à présent enrichi des connecteurs extensionnels habituels, $\wedge$ et $\vee$. Une formule *de degré zéro* (zdf) est une formule qui ne contient que les connecteurs $\wedge$, $\vee$ et $\neg$; une *conséquence du premier degré* (ou fde, « *first degree entailment* ») est une formule de la forme $(A \rightarrow B)$, où A et B sont des formules de degré zéro. Le système $\mathbf{R}_{\mathbf{fde}}$ (ou $\mathbf{E}_{\mathbf{fde}}$; les deux systèmes sont identiques) ne concerne que les conséquences du premier degré, à l'exclusion des formules où des occurrences de $\rightarrow$ figurent dans la portée de ce même connecteur (emboîtements de conséquences), ou dans la portée des connecteurs extensionnels. On peut donc interpréter à volonté $\rightarrow$, dans le contexte de ce système, comme un connecteur d'implication, ou comme signifiant une relation métalinguistique de conséquence logique entre expressions vérifonctionnelles : c'est donc ici que les comparaisons possibles avec les systèmes classiques sont les plus directes.

On définit d'abord une notion de conséquence du premier degré valide, la conséquence tautologique, qui est ultérieurement représentée axiomatiquement. Le syllogisme disjonctif (SD) n'est pas valide au sens de cette notion : on examine brièvement les arguments fournis par Anderson et Belnap contre le SD. On présente ensuite une sémantique algébrique pour la conséquence du premier degré. Il va sans dire que, à elle seule, cette construction algébrique, bien qu'elle invalide le SD, ne résout nullement la question de sa valeur en tant que mode d'inférence admissible.

LA NOTION DE CONSÉQUENCE TAUTOLOGIQUE

Le premier problème de ce chapitre est de trouver des critères plausibles permettant d'épingler, parmi les conséquences du premier degré, celles qui sont valides au sens de la conséquence pertinente, qu'on appellera *conséquences tautologiques.* Il est d'emblée clair que la relation d'implication logique classique ne peut être retenue à titre de critère. En effet, une contradiction comme $(A \wedge \neg A)$ implique classiquement n'importe quelle formule B, alors qu'une contradiction est en général sans pertinence pour une proposition B quelconque (c'est pour des raisons analogues, de faute de pertinence, que $\neg A \rightarrow (A \rightarrow B)$ a été refusé au chapitre précédent). Cependant, on peut admettre que $(A \wedge \neg A)$ entraîne A, puisque ce n'est qu'un cas particulier du principe plausible lié à la conjonction : la conjonction $A \wedge B$ entraîne A (et B également)[1]. Il faut donc trouver un moyen systématique de départager ces cas.

On isole d'abord à cette fin les *conséquences primitives.* Appelons dans ce contexte *atomes* les variables propositionnelles et leurs négations (une formule de la forme $\neg p$ est donc un atome, comme p elle-même).

DÉFINITION 1. — Une *conséquence primitive* est une formule où l'antécédent est une *conjonction primitive* (*i.e.* une conjonction d'atomes), le conséquent une *disjonction primitive* (*i.e.* une disjonction d'atomes), *i.e.* une formule de la forme :

$$A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m,$$

où chaque A_i et B_j est un atome (il n'est pas exclu que n, ou m, ou les deux, = 1).

DÉFINITION 2. — Une conséquence primitive est valide, ou *explicitement tautologique,* ssi quelque atome A_i est identique à quelque atome B_j.

Exemples 1 :

Les formules : $(p \rightarrow p)$, $(p \wedge q) \rightarrow q$, $p \rightarrow (p \vee q)$ sont explicitement tautologiques[2].

Le sont de même les formules :

$(\neg p \rightarrow \neg p)$, $(p \wedge \neg p) \rightarrow \neg p$, $q \rightarrow (q \vee \neg q)$, $(p \wedge q) \rightarrow (r \vee q)$.

En revanche, ne sont pas explicitement tautologiques, et sont donc invalides, les conséquences primitives suivantes :

$(p \wedge \neg p) \rightarrow q$, $p \rightarrow (q \vee \neg q)$, $(p \wedge \neg p) \rightarrow (q \vee \neg q)$, etc.

Bien sûr, toutes les conséquences du premier degré ne sont pas des conséquences primitives, par exemple parce que le conséquent est une conjonction. Que penser ainsi de la formule :

$p \rightarrow (p \wedge (r \vee \neg r))$?

La réponse naturelle est qu'elle ne serait acceptable que si les deux conséquences primitives $p \rightarrow p$ et $p \rightarrow (r \vee \neg r)$ étaient explicitement tautologiques, ce qui n'est pas le

1. Le « connexivisme » refuse le principe de Simplification, A et B entraîne A, au motif que, dans le cas particulier où B est non-A, la négation annule la position antérieure de A Je ne poursuivrai pas cette voie.

2. L'implication *analytique*, proposée par W. Parry au début des années 1930, est fondée sur le refus de la dernière formule, au motif qu'une variable y figure dans le conséquent qui ne figure pas dans l'antécédent. L'idée de Parry était de donner sens à la définition kantienne de l'analyticité (le prédicat est contenu dans le sujet), en exigeant d'une implication valable que le contenu du conséquent soit contenu dans le contenu de l'antécédent (principe du *content containment).* Il faut alors rejeter Contraposition, au motif que, s'il y a moins de variables dans B que dans A, il y aura plus de variables dans le conséquent de $\neg B \rightarrow \neg A$ que dans l'antécédent. Le système de l'implication analytique diverge ici de la logique pertinente (voir Parry, 1989).

cas. Pourquoi « naturelle » ? Si l'on acceptait comme valides les formules de la forme $A \rightarrow (A \wedge (B \vee \neg B))$, on se trouverait dans la situation suivante : comme la conjonction $A \wedge (B \wedge \neg B)$ implique chacun de ses conjoints, donc en particulier $(B \vee \neg B)$, on obtiendrait par transitivité $A \rightarrow (B \vee \neg B)$, dont justement on ne veut pas. Une solution serait de renoncer à la transitivité de la conséquence, mais ce serait certainement une solution désespérée qu'Anderson et Belnap récusent en ces termes :

> « Tout critère selon lequel la conséquence n'est pas transitive est *ipso facto* incorrect. Il paraît en fait incroyable que quelqu'un puisse admettre que B découle de A, que C découle de B, mais pense que quelque argument supplémentaire est requis pour établir que A entraîne C. Quelle meilleure justification de $A \rightarrow C$ pourrait-on trouver ? » (Anderson et Belnap, 1975, § 15).

La formule $A \rightarrow (A \wedge (B \vee \neg B))$ doit donc être refusée. Cette remarque invite à proposer le critère suivant (précédé d'une définition préliminaire), pour les formules qui ne sont pas des conséquences primitives.

DÉFINITION 3. — Une conséquence du premier degré est en *forme normale* <u>ssi</u> l'antécédent est une disjonction de conjonctions primitives, et le conséquent une conjonction de disjonctions primitives, *i.e.* a la forme :

$$A_1 \vee \ldots \vee A_n \rightarrow B_1 \wedge \ldots \wedge B_m,$$

où chaque A_i est une conjonction primitive, et chaque B_j est une disjonction primitive.

Exemples 2 :

$p \rightarrow (p \wedge (q \vee \neg q))$, $((p \vee \neg p) \vee q) \rightarrow q$, $(p \vee q) \rightarrow (p \wedge q)$ sont en forme normale ; de même :

$$\neg p \vee (p \wedge \mathrm{q}) \rightarrow (\neg p \vee p) \wedge (\neg p \vee q).$$

DÉFINITION 4. — Une conséquence en forme normale :

$$A_1 \vee \ldots \vee A_n \rightarrow B_1 \wedge \ldots \wedge B_m,$$

est *explicitement tautologique* <u>ssi</u> pour tout A_i et tout B_j la conséquence primitive $A_i \rightarrow B_j$ est explicitement tautologique.

Exemples 3 :

On vérifiera facilement que, dans les exemples 2, les formules de la première ligne ne sont pas valides (faute d'être explicitement tautologiques) ; en revanche, la formule de la seconde ligne l'est (une forme de distributivité de $\vee$ par $\wedge$).

Ces définitions ne recouvrent pas encore tous les cas possibles, puisque toutes les conséquences du premier degré ne sont pas en forme normale au sens défini (exemple : $(p \wedge (\neg p \vee q)) \rightarrow q$ n'est pas en forme normale, l'antécédent n'étant pas une disjonction de conjonctions primitives). On utilise donc les règles usuelles de remplacement (associativité, commutativité, suppression des doubles négations, distributivité et lois de De Morgan de rentrée de la négation), qui permettent de transformer une formule vérifonctionnelle en une formule « équivalente » en forme normale, disjonctive (pour les antécédents), et conjonctive (pour les conséquents). On peut admettre en effet que ces règles de transformation préservent la validité, *i.e.* qu'une conséquence $A \rightarrow B$ est valide si et seulement si, après transformation, une forme normale obtenue l'est aussi.

Toute conséquence du premier degré peut donc être transformée, par application itérée de ces règles, en une conséquence en forme normale de la forme appropriée.

DÉFINITION 5. - Une conséquence du premier degré $A \rightarrow B$ est une *conséquence tautologique* ssi elle a une forme normale $A_1 \vee \ldots \vee A_n \rightarrow B_1 \wedge \ldots \wedge B_m$ explicitement tautologique.

La notion ainsi définie peut être comprise comme une condition nécessaire et suffisante pour la validité d'une conséquence entre fonctions de vérité. Notons qu'il s'agit d'un critère purement syntaxique de validité. Elle est évidemment *décidable*[1].

Exemple 4 :

Soit la formule :

$\big((\neg p \vee q) \wedge (\neg q \vee r)\big) \rightarrow (\neg p \vee r)$;

une forme normale en est :

$(\neg p \wedge \neg q) \vee (q \wedge \neg q) \vee (r \wedge \neg p) \vee (r \wedge q) \rightarrow (\neg p \vee r)$,

qui n'est pas explicitement tautologique : voir la conséquence $(q \wedge \neg q) \rightarrow (\neg p \vee r)$, conséquence primitive non explicitement tautologique. Or la formule de départ peut être réécrite :

$$\big((p \supset q) \wedge (q \supset r)\big) \rightarrow (p \supset r),$$

avec le signe pour le conditionnel matériel pris comme abréviation ($A \supset B$ abrège $\neg A \vee B$). Le syllogisme conjonctif avec le conditionnel matériel n'est donc pas valide : la conjonction des deux prémisses n'entraîne pas la conclusion.

LE SYLLOGISME DISJONCTIF

Les remarques qui suivent peuvent être rangées sous la rubrique habituelle « fautes de pertinence » ; mais, cette fois, elles sont à première vue déconcertantes.

La conséquence la plus frappante du critère de validité ainsi proposé pour les conséquences du premier degré est qu'il exclut la formule :

$$\big(\neg A \wedge (A \vee B)\big) \rightarrow B$$

qui doit être réputée invalide (une instance comme : $\neg p \wedge (p \vee q) \rightarrow q$, après mise en forme normale, *n'est pas* explicitement tautologique). Puisque $\neg A \wedge (A \vee B)$ n'entraîne pas B, au sens de la conséquence tautologique, on est donc conduit à rejeter la règle d'inférence : de $\neg A$ et de $(A \vee B)$, inférer B, *i.e.* la règle du syllogisme disjonctif (SD), comme principe d'inférence. Or abandonner ce principe, qui est littéralement le cinquième indémontrable Stoïcien, n'a rien d'évident. Comme le dit Michael Dunn de manière humoristique :

> « Ce rejet est typiquement la chose la plus difficile à faire avaler en ce qui concerne la logique pertinente. On commence par quelques motivations fort plaisantes à propos de l'implication pertinente, et de l'usage des indices souscrits pour garder trace de l'utilisation

1. C'est la dernière occasion que j'aurai d'évoquer les questions de décidabilité. Pour mémoire, les fragments implication, et implication + négation sont décidables, le système **R** ne l'est pas (Urquhart, 1984).

réelle d'une hypothèse, et on arrive au point où l'on dit "et bien sûr nous abandonnons le syllogisme disjonctif" ; et là on perd son public » (Dunn, 1986).

On le conçoit : si le connecteur ⊃ est introduit comme abréviation, le rejet du syllogisme disjonctif devient le rejet de la règle de Détachement (c'est-à-dire du Modus Ponens) pour le conditionnel matériel, ici sous la forme :

de $\neg A$ et de $(\neg A \supset B)$, inférer B.

Une bonne façon d'aborder la question du syllogisme disjonctif est de reprendre la « preuve » de Lewis et Langford, 1932, selon laquelle, en dépit des considérations déjà évoquées, A & non-A entraîne B, c'est-à-dire la preuve du principe classique *Ex Falso Quodlibet Sequitur* (EFQ). On défend d'ordinaire la validité de ce principe en invoquant l'idée selon laquelle l'essence de la relation de conséquence réside dans la préservation de la vérité : si on admet que la conjonction A & non-A n'est pas vraie et ne saurait l'être (ce qui par ailleurs peut être refusé ; mais laissons ce point de côté pour l'instant), on peut dire que ce n'est jamais le cas que la prémisse A & non-A est vraie, et que B est faux, pour n'importe quel B. Donc une contradiction entraîne n'importe quoi. La preuve de Lewis et Langford, quant à elle, est une tentative pour montrer pas à pas que EFQ découle de principes d'inférence que tout le monde accepte comme valides. Or au nombre de ces principes figure justement le Syllogisme disjonctif.

En fait, l'argument de Lewis et Langford était déjà connu des logiciens médiévaux (même s'il a peut-être été retrouvé indépendamment). D'après William et Martha Kneale, dans les *In Universam Logicam Quaestiones*, attribuées à Duns Scot, puis au Pseudo-Scot, on trouve la preuve suivante (je traduis librement). De *Socrate existe et Socrate n'existe pas*, on peut inférer *Socrate n'existe pas* ; et pour la même raison, *Socrate existe*. De cette dernière proposition, s'ensuit *Socrate existe, ou l'homme est un âne*. On a donc : *ou Socrate existe, ou l'homme est un âne* ; mais *Socrate n'existe pas*. Donc *l'homme est un âne*. Le Pseudo-Scot ajoute que chaque pas inférentiel est « formellement valide » (« *omnes consequentiae sunt formales* »), *i.e.* valide selon la forme seule et non en raison de la signification des termes qui y figurent[1].

L'argument de Lewis et Langford, 1932, est essentiellement le même. Supposons qu'on admette les principes d'inférence suivants :

1) de A & B, on peut inférer A, et inférer B	Simplification
2) de A, on peut inférer A ou B	ou-Introduction
3) de A ou B, et de non-A, on peut inférer B	SD

Ces principes admis, la preuve que A & non-A entraîne B est la suivante :

a) A &non-A	prémisse
b) A	par Simplification
c) non-A	par Simplification
d) A ou B	par ou-Introduction sur *b*)
e) B	par SD

Si l'on accepte les trois principes d'inférence proposés, ainsi que la transitivité de la conséquence $(A \wedge \neg A) \rightarrow B)$ doit être acceptée comme une conséquence valide, puisque

1. Kneale, 1962, p. 281.

B peut être déduit de la prémisse contradictoire. Comme les autres principes, qui ne sont que l'élimination de la conjonction, et l'introduction de la disjonction, semblent innocents, et que la conclusion $(A \wedge \neg A) \rightarrow B)$ commet une faute caractéristique de pertinence, on est donc amené, pour dynamiter la preuve, à rejeter comme fautif le principe du syllogisme disjonctif (qui par ailleurs, on vient de le voir, ne satisfait pas le critère *formel* pour la conséquence tautologique). Il y a ici, il faut l'avouer, un air de paradoxe : n'est-il pas correct, de la disjonction A ou B, d'inférer B si l'on sait que A n'est pas le cas ? Si la disjonction est vraie, dira-t-on, alors que l'un de ses membres est faux, c'est que l'autre est vrai.

Qu'est-ce donc qui ne va pas, avec le syllogisme disjonctif[1] ? Une *première* tentative de réponse est en fait insuffisante (il faut bien reconnaître qu'Anderson et Belnap, 1975, n'est pas satisfaisant sur ce point).

Face à l'incrédulité que provoque le rejet de SD, on explique cette réaction, pour l'atténuer, en faisant remarquer qu'il y a en effet des cas où le syllogisme disjonctif est bien valide. C'est seulement une généralisation abusive, fondée sur une équivoque concernant la disjonction, qui nous fait croire qu'il est *toujours* valide. Le syllogisme disjonctif est acceptable, et probablement réellement utilisé dans nos raisonnements, quand le « ou » de « A ou B » a une signification intensionnelle, quand il veut dire quelque chose comme « si non-A, alors B ». Il y a, en effet, des raisonnements à partir de non-A, et de A ou B, parfaitement corrects, précisément parce qu'il y a une relation de pertinence entre A et B. Par exemple, de :

> ou nous arrivons à l'heure à la gare, ou nous ratons le train,

et de :

> nous n'arriverons pas à l'heure,

il est correct d'inférer que nous allons rater le train. Mais ici le « ou » de la première prémisse n'est pas vérifonctionnel ; le lien entre A et B est assuré par le fait que précisément non-A entraîne B : si nous n'arrivons pas à l'heure, nous raterons le train (comparer avec : « Ou nous arrivons à l'heure, ou il y aura une éclipse de soleil »).

Un argument général est susceptible de confirmer cette remarque : si, sachant que A n'est pas le cas (est faux), on affirme cependant la disjonction A ou B, ce ne peut être pour des raisons vérifonctionnelles (*i.e.* parce que l'un des deux constituants est connu comme vrai - à moins qu'on ne sache déjà que B est vrai, auquel cas l'inférence est purement rhétorique). Car A étant faux ou connu comme tel, on ne peut prétendre inférer la disjonction A ou B à partir de A. Dans les termes de Stephen Read :

> « Si la justification pour asserter "P ou Q" est, disons, P, apprendre que P est faux (*i.e.* que non-P est vrai), loin de nous permettre d'inférer Q, supprime la garantie pour asserter "P ou Q" en premier lieu » (Read, 1988).

Donc le « ou » de la disjonction dans les raisonnements corrects qui ont la forme du SD n'est pas purement extensionnel, et il y a une faute à le représenter par $\vee$. Dans de tels raisonnements, si l'on pose A ou B, ce n'est pas en vertu de la vérité connue ou supposée de A. C'est au contraire parce que non-A entraîne déjà à lui seul B : comme on

1. Pour reprendre le titre d'un article de Read, « What is wrong with Disjunctive Syllogism ? » (Read, 1981).

a non-A, par Modus Ponens, on peut poser B. On voit alors que la « preuve » de Lewis est fondée sur une équivoque concernant « ou ». Le passage de A à A ou B n'est justifié que si on prend « ou » de manière purement vérifonctionnelle, et si on représente la conclusion par $A \vee B$, car au sens intensionnel de « ou », cette inférence n'est pas valide. En revanche, le SD utilisé à la dernière ligne n'est valide que si « ou » est compris intensionnellement. Voilà levée l'équivoque sur laquelle notre malaise était fondé.

> « En rejetant le principe du syllogisme disjonctif, nous ne voulons le restreindre qu'aux cas où le "ou" est pris de manière vérifonctionnelle. En général, et en ce qui concerne nos raisonnements ordinaires, ce n'est pas le cas ; il est possible que quand le principe est utilisé en raisonnant, on ait en tête un sens intensionnel de "ou", où il y a une relation de pertinence entre les disjoints. Mais pour le sens intensionnel de "ou" il est clair que l'analogue de $A \rightarrow (A \vee B)$ est invalide, puisque cela ne vaudrait que si la simple vérité de A était suffisante pour la pertinence de A relativement à B ; il y a donc un sens selon lequel la faille réelle dans l'argument de Lewis n'est pas une faute de pertinence, mais une faute d'ambiguïté » (Anderson et Belnap, 1975, § 16.1).

Tout cela est bel et bon. L'impression de validité provient d'un usage intensionnel de « ou » dans les syllogismes disjonctifs corrects, abusivement confondu avec la disjonction vériconditionnelle. La clarification proposée a de plus le mérite de mettre le doigt sur un nouveau connecteur, la disjonction intensionnelle, baptisée ultérieurement « fission », qu'on pourra définir ainsi dans **R** :

$$A \oplus B =_{\text{Def}} \neg A \rightarrow B,$$

et dont il sera fait usage plus tard (voir chap. 7). Mais il faut bien avouer que ce faisant, on n'a pas clairement répondu à la question : pourquoi le SD, le « bon vieux » SD avec sa disjonction extensionnelle (et sa négation également supposée telle), n'est-il pas valide ? Si $A \vee B$ est vrai, l'un des deux disjoints l'est également ; comme A est faux, B doit être vrai, *point à la ligne.* C'est ce que Read a baptisé le « défi de Geach »[1]. On verra, aux chapitres 5 et 6 en particulier, les conséquences des efforts faits pour y répondre[2].

1. Read, 1988, chap. 7, à propos du compte rendu par Geach d'Anderson et Belnap, 1975 (Geach, 1977).

2. On ne confondra pas le refus de l'inférence : de $\neg A$ et de $A \vee B$, inférer B, ou le rejet du syllogisme disjonctif, avec la question de l'admissibilité de la règle dite « règle γ d'Ackermann » : si $\vdash A$ et si $\vdash \neg A \vee B$, alors $\vdash B$, *i.e.* si A et $\neg A \vee B$ sont des théorèmes, alors B est aussi un théorème. Une règle d'inférence peut être invalide, et cependant applicable quand ses prémisses sont des théorèmes, auquel cas la conclusion l'est aussi (voir la règle de nécessitation en logique modale). C'est le cas de la règle en question, qui est admissible pour **R**, au sens où elle n'ajoute rien au stock des théorèmes. La question ne sera pas poursuivie ici.

LE SYSTÈME $\mathbf{R_{fde}}$ POUR LA CONSÉQUENCE TAUTOLOGIQUE

Le système suivant capture précisément la notion définie plus haut de conséquence tautologique :

Axiomes :

1)	$A \wedge B \to A$	$A \wedge B \to B$	$\wedge$Élimination
2)	$A \to A \vee B$	$B \to A \vee B$	$\vee$Introduction
3)	$A \wedge (B \vee C) \to (A \wedge B) \vee C$		Distributivité
4)	$A \to \neg\neg A$	$\neg\neg A \to A$	Double négation

Les règles d'inférence sont nombreuses, faute de pouvoir être exprimées sous forme d'axiomes en raison des limitations sur la formation des formules (pas de $\to$ dans la portée de connecteurs extensionnels) :

5) de $A \to B$ et $B \to C$, inférer $A \to C$	Transitivité
6) de $A \to B$ et $A \to C$, inférer $A \to (B \wedge C)$	$\wedge$Introduction
7) de $A \to C$ et $B \to C$, inférer $(A \vee B) \to C$	$\vee$Introduction
8) de $A \to B$, inférer $\neg B \to \neg A$	Contraposition

Pour montrer que $\mathbf{R_{fde}}$ axiomatise adéquatement la classe des conséquences tautologiques telle que définie plus haut, il faut démontrer deux propositions : la première énonce une forme de complétude (toutes les conséquences tautologiques sont prouvables), la seconde une forme de correction (fiabilité), relativement bien sûr à la notion de validité définie.

PROPOSITION 1. — Toutes les conséquences tautologiques sont des théorèmes de $\mathbf{R_{fde}}$.

(i) Toutes les conséquences primitives valides sont prouvables (noter que les axiomes pour la double négation, plus transitivité, assurent que $A \to A$ est prouvable), par les règles pour la conjonction et la disjonction.

(ii) Les conséquences tautologiques en forme normale sont prouvables ;

Exemple 5 :

On considère la formule en forme normale $(p \wedge \neg q) \vee (r \wedge \neg s) \to (p \vee s) \wedge (\neg q \vee s)$, qui est bien une conséquence tautologique. $(p \wedge \neg q) \to (p \vee s)$ est explicitement tautologique ; de même $(p \wedge \neg q) \to (\neg q \vee s)$; donc tous deux sont des théorèmes. Par la règle 6) $(p \wedge \neg q) \to (p \vee s) \wedge (\neg q \vee s)$ est aussi un théorème. Même raisonnement pour $(r \wedge \neg s) \to (p \vee s) \wedge (\neg q \vee s)$. Par 7) $(p \wedge \neg q) \vee (r \wedge \neg s) \to (p \text{ vs}) \wedge (\neg q \text{ vs})$, est aussi un théorème.

En général, si on a une conséquence en forme normale $A_1 \vee ... \vee A_n \to B1 \wedge ... \wedge B_m$ explicitement tautologique, pour tout A_i et tout B_j, la conséquence primitive $A_i \to B_j$ est explicitement tautologique, donc prouvable par (i). Il s'ensuit par la règle 6) que :

$$A_l \to B_1 \wedge \ldots \wedge B_m$$

—

—

$$A_n \to B_1 \wedge ... \wedge B_m$$

sont prouvables ; d'où par la règle 7), $A_1 \vee \ldots \vee A_n \rightarrow B_1 \wedge \ldots \wedge B_m$ est un théorème.

(iii) Le dernier point consiste à vérifier qu'une conséquence tautologique qui n'est pas en forme normale est prouvable. Puisqu'il s'agit d'une conséquence tautologique, elle a une forme normale explicitement tautologique, qui donc est prouvable ; il faut donc montrer qu'on peut» remonter » à la formule de départ, et la prouver à partir de la forme normale. Pour cela, il faut s'assurer :

– qu'aux règles de transformation mentionnées plus haut correspondent des théorèmes de $\mathbf{R_{fde}}$, par exemple à la règle de Commutativité, correspondent les théorèmes de co-implication :

$(A \wedge B) \rightarrow (B \wedge A)$ et $(B \wedge A) \rightarrow (A \wedge B)$ (abrégés en $A \wedge B \leftrightarrow B \wedge A$). On a également le théorème : $A \wedge (B \vee C) \leftrightarrow (A \wedge B) \vee (A \wedge C)$

- dans un sens :

$A \wedge (B \vee C) \rightarrow A$	
$A \wedge (B \vee C) \rightarrow (A \wedge B) \vee C$	Distributivité

d'où :

$A \wedge (B \vee C) \rightarrow A \wedge \big((A \wedge B) \vee C\big)$	$\wedge$Introduction
$A \wedge (B \vee C) \rightarrow A \wedge \big(C \vee (A \wedge B)\big)$	Commutativité
$A \wedge (B \vee C) \rightarrow (A \wedge C) \vee (A \wedge B)$	Distributivité, Transitivité
$A \wedge (B \vee C) \rightarrow (A \wedge B) \vee (A \wedge C)$	Commutativité

- dans l'autre, par application des règles pour la conjonction et la disjonction. Les quatre lois de De Morgan sont également démontrables dans $\mathbf{R_{fde}}$, y compris donc la forme classique : $\neg(A \wedge B) \rightarrow \neg A \vee \neg B$ (utiliser Contraposition classique deux fois).

– enfin, qu'on a un théorème de remplacement usuel disant que si A co-implique B, il en est de même de F(A) et F(B), où F(B) est obtenu à partir de F(A) par remplacement d'une ou plusieurs occurrences de A par B.

Il découle de ces considérations que toute conséquence tautologique est prouvable.

PROPOSITION 2. — Tous les théorèmes de $\mathbf{R_{fde}}$ sont des conséquences tautologiques.

Il faut d'abord montrer que tous les axiomes sont des conséquences tautologiques (par récurrence sur la complexité des formules) et que les règles préservent cette propriété (ce qui est clair en ce qui concerne les règles pour la conjonction et la disjonction).

Pour la règle Contraposition, il faut montrer que si $A \rightarrow B$ est une conséquence tautologique, $\neg B \rightarrow \neg A$ l'est également. Or, étant donné une forme normale disjonctive d'une formule C, on obtient une forme conjonctive de $\neg C$ en échangeant les connecteurs $\wedge$ et $\vee$, et en remplaçant les variables par leurs négations et inversement. Donc si une forme normale disjonctive de A a la propriété requise (*i.e.* chaque $A_i \rightarrow B_j$ est explicitement tautologique) avec une forme conjonctive de B, une forme disjonctive de $\neg B$ l'aura aussi avec une forme conjonctive de $\neg A$.

Pour la règle Transitivité : supposons que $A \rightarrow B$, $B \rightarrow C$ soient des conséquences tautologiques, qu'on peut supposer mises en formes normales :

$A_1 \vee ... \vee A_n \rightarrow B_1 \wedge ... \wedge B_m$
$B_1' \vee ... \vee B_b' \rightarrow C_1 \wedge ... \wedge C_p$

A_1 partage un atome avec chaque B_j ; soit $P_1 \vee ... \vee P_n$ la conjonction de ces atomes. $P_1 \vee ... \vee P_n$ est l'un des disjoints de $B_1' \vee ... \vee B_b'$, puisque cette disjonction est faite de conjonctions d'atomes pris dans chacun des $B_1 \wedge ... \wedge B_m$ de toutes les manières possibles (distribution « croisée »). Comme par hypothèse chaque B_i' a un atome en commun avec chaque C_j , $P_1 \vee ... \vee P_n$, et donc A_1, a un atome en commun avec chaque C_j. Le même raisonnement vaut pour les autres disjoints de $A_1 \vee ... \vee A_n$. Donc $A_1 \vee ... \vee A_n \rightarrow C_1 \wedge ... \wedge C_p$, et donc aussi A → C, sont des conséquences tautologiques.

UNE SÉMANTIQUE ALGÉBRIQUE POUR $\mathbf{R_{fde}}$

L'objet de ce paragraphe est de présenter, puis d'utiliser, un type de structure algébrique, les *treillis de De Morgan,* qui puisse être considéré comme fournissant au langage pour $\mathbf{R_{fde}}$ une interprétation. Ces structures seront au calcul de la conséquence du premier degré ce que les algèbres de Boole sont au calcul propositionnel classique.

La notion de treillis de De Morgan[1]

Un treillis T est un ensemble (non vide) muni de deux opérations ∧ et ∨, satisfaisant les propriétés suivantes :

$a \wedge a = a$	$a \vee a = a$	Idempotence
$a \wedge b = b \wedge a$	$a \vee b = b \vee a$	Commutativité
$a \wedge (b \wedge c) = (a \wedge b) \wedge c$	$a \vee (b \vee c) = (a \vee b) \vee c$	Associativité
$a \wedge (a \vee b) = a$	$a \vee (a \wedge b) = a$	Absorption

Une relation d'ordre partiel sur T, *i.e.* réflexive, transitive et antisymétrique, est définie par :

$a \leq b$ si et seulement si $a \wedge b = a$

($a \wedge b$ est appelé la « borne inférieure » de a et de b, ou « le plus grand des plus petits » selon la relation ≤ ; et $a \vee b$ la « borne supérieure » de a et de b, ou « le plus petit des plus grands »).

Un treillis *distributif* est un treillis satisfaisant les deux lois suivantes (qui s'impliquent mutuellement) :

$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ $\qquad$ $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$.

Certains sous-treillis de treillis sont importants ; un *filtre* d'un treillis T est un sous-ensemble (non vide) F de T tel que :

(l) si $a, b \in$ F, alors $a \wedge b \in$ F ;
(2) si $a \in$ F, alors $a \vee b \in$ F (ce qui implique que si $a \in$ F, et $a \leq b$, $b \in$ F).

Un filtre est *premier* s'il satisfait en outre :

1. On parle aussi d'algèbres quasi booléennes.

(3) si $a \vee$ b $\in$ F, alors ou $a \in$ F, ou b $\in$ F.

Un treillis distributif est un *treillis de De Morgan* s'il satisfait en outre, pour une opération unaire -, parfois dite « complémentation de De Morgan », les propriétés suivantes :

$a \leq --a$, $--a \leq a$	Double négation classique
si $a \leq b$, alors $-b \leq -a$, si $-a \leq -b$, alors $b \leq a$	Contraposition classique

Le nom « de De Morgan » est justifié par le fait que ces propriétés permettent de montrer qu'un tel treillis satisfait les lois de De Morgan, sous la forme :

(l) $-(a \wedge b) = -a \vee -b$ (2) $-(a \vee b) = -a \wedge -b$.

Montrons (l) à titre d'exemple :

Comme $a \wedge b \leq a$, et $a \wedge b \leq b$, on a : $-a \leq -(a \wedge b)$, et $-b \leq (a \wedge b)$, par Contraposition. Donc $-a \vee -b$ (le « plus petit des plus grands ») $\leq -(a \wedge b)$.

Par ailleurs, comme $-a \leq -a \vee -b$, et $-b \leq -a \vee -b$, on a $-(-a \vee -b) \leq --a$, et également$-(-a \vee -b) \leq --b$; d'où$-(-a \vee -b) \leq a$ et $-(-a \vee -b) \leq b$, par Double négation classique. Donc $-(-a \vee -b) \leq a \wedge b$ (« le plus grand des plus petits »). Par contraposition, on obtient $-(a \wedge b) \leq -a \vee -b$. Finalement, $-(a \wedge b) = -a \vee -b$.

(2) se démontre par un raisonnement analogue.

Un treillis de De Morgan reflète nombre des propriétés de la négation classique, mais les deux propriétés suivantes lui font défaut :

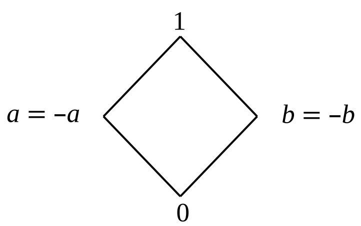
$a \wedge -a \leq b$, pour tout élément b,

$a \leq b \vee -b$, pour tout élément a.

Si ces deux propriétés sont rajoutées, un treillis de De Morgan devient une *algèbre de Boole, i.e.* un treillis distributif complémenté avec un plus petit élément $0 = a \wedge -a$, et un plus grand élément $1 = a \vee -a$.

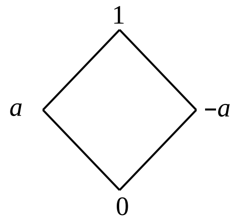
Les deux diagrammes suivants représentent respectivement : 1) un treillis de De Morgan à quatre éléments, le treillis **4** ; 2) une algèbre de Boole avec le même nombre d'éléments. On observera que le premier possède une propriété remarquable : *a* et *b* sont identiques à leurs images par l'opération $-$ (autrement dit, l'opération $-$ a des points fixes).

1) Le treillis **4** (Dunn, 1986)[1] :



1. Le § 18 d'Anderson et Belnap, 1975, consacré aux Algèbres intensionnelles et rédigé par M. Dunn, caractérise $R_{\mathbf{fde}}$ par un treillis intensionnel à 8 éléments (voir ci-dessous p. 67).

2) Une algèbre de Boole à 4 éléments :



Une dernière définition : Anderson et Belnap, 1975 (§ 18), privilégie, en rapport avec le système $\mathbf{R}_{\mathbf{fde}}$, les *treillis de De Morgan intensionnels.* Un treillis de De Morgan intensionnel est une structure algébrique $\langle T, \bigwedge, \bigvee, -\rangle$ où l'opération complémentation n'a pas de point fixe, *i.e.* pour aucun élément a, $a = -a$ (la désignation « intensionnel » sera justifiée plus loin). Un filtre F de T est *consistant* si pour aucun a élément de T, on n'a à la fois $a \in F$ et $-a \in F$, et *complet* si pour tout a de T, ou $a \in F$, ou $-a \in F$. La condition : pour aucun élément a, $a = -a$, est équivalente à l'existence d'un filtre consistant et complet (non démontré ici).

On peut observer que le diagramme 1) ci-dessus n'est pas un treillis intensionnel, la complémentation de De Morgan ayant des points fixes (a et b sont identiques à leurs « complémentaires »). Le sous-ensemble $\{a, 1\}$, avec $a \leq 1$, est bien un filtre (non le seul), mais on peut vérifier qu'aucun filtre n'est à la fois consistant et complet : Le diagramme 2), en revanche, est bien un treillis intensionnel, et en fait une algèbre de Boole.

Les treillis de De Morgan comme modèles pour $\mathbf{R}_{\mathbf{fde}}$

Un modèle M pour les conséquences du premier degré est une paire (T, v), où T est un treillis de De Morgan (non nécessairement intensionnel), et v une valuation (de degré zéro), *i.e.* une application des formules de degré zéro (uniquement) dans T, satisfaisant :

$$v(\neg A) = -v(A)$$
$$v(A \wedge B) = v(A) \bigwedge v(B)$$
$$v(A \vee B) = v(A) \bigvee v(B)$$

On dit qu'une conséquence du premier degré $A \rightarrow B$ est *vraie dans un modèle* $M = \langle T, v\rangle$ si et seulement si $v(A) \leq v(B)$ *;* $A \rightarrow B$ est *valide dans un treillis* T si et seulement si elle est vraie pour toute valuation dans T ; enfin $A \rightarrow B$ est *valide* si et seulement si elle est valide dans tout treillis, et sinon *falsifiable.*

On peut à présent montrer qu'une conséquence du premier degré $A \rightarrow B$ est valide si et seulement si $A \rightarrow B$ est un théorème de $\mathbf{R}_{\mathbf{fde}}$. Auparavant, un dernier détour algébrique est nécessaire, afin de définir la notion de *modèle canonique.*

L'Algèbre de Lindenbaum de $\mathbf{R}_{\mathbf{fde}}$

La relation de co-implication entre deux formules A et B a lieu quand A et B sont interdéductibles, *i.e.* quand $A \rightarrow B$ et $B \rightarrow A$ sont des théorèmes. La relation de co-implication est évidemment une relation d'équivalence sur les formules. On peut donc opérer une partition des formules de degré zéro en classes d'équivalence pour cette relation, la classe d'équivalence d'une formule A étant notée [A].

On peut alors définir des opérations sur ces classes d'équivalence en posant :

$- [A] = [\neg A]$
$[A] \wedge [B] = [A \wedge B]$
$[A] \vee [B] = [A \vee B]$.

Le théorème de remplacement assure que les opérations sont bien définies, au sens où elles ne dépendent pas du choix du représentant. On peut de plus définir une relation d'ordre entre classes d'équivalence :

$[A] \leq [B]$ si et seulement si $A \rightarrow B$ est un théorème de $\mathbf{R_{fde}}$.

Il est aisé de voir que l'Algèbre de Lindenbaum pour $\mathbf{R_{fde}}$, ainsi définie, est un treillis de De Morgan (voir la formulation des axiomes). On définit alors le *modèle canonique* à partir de cette algèbre en spécifiant la valuation canonique **c** de la manière suivante :

pour toute formule A, $\mathbf{c}(A) = [A]$.

On peut alors démontrer les deux théorèmes attendus : le système $\mathbf{R_{fde}}$ est correct pour la validité, il est d'autre part complet pour cette même notion.

THÉORÈME 1. — Si une conséquence du premier degré est prouvable dans $\mathbf{R_{fde}}$, elle est valide.

Les axiomes sont valides ; par exemple, l'axiome $(A \wedge B) \rightarrow A$ est manifestement valide, car dans tout treillis, $v(A) \wedge v(B) \leq v(A)$, quelle que soit *v*. Et les règles préservent la validité.

THÉORÈME 2. — Si une conséquence du premier degré est valide, elle est prouvable.

Supposons qu'une formule $A \rightarrow B$ ne soit pas prouvable. Par la définition de la relation d'ordre partiel $\leq$ sur les classes d'équivalence, *on n'a pas* $[A] \leq [B]$, *i.e.* $\mathbf{c}(A)$ n'est pas inférieur ou égal à $\mathbf{c}(B)$. La formule n'est donc pas vraie dans le modèle canonique, elle est donc falsifiable, *i.e.* non valide.

On a donc un théorème de complétude pour $\mathbf{R_{fde}}$. Il ne faut pas exagérer, cependant, la valeur de ce résultat, comme si la notion de validité en jeu jouissait d'une stabilité et d'une évidence intrinsèque capables de justifier la notion formelle de prouvabilité. La notion de validité dans tout treillis de De Morgan a quelque chose de manifestement *ad hoc*, et ne donne qu'une représentation algébrique de formules dont l'acceptabilité vient d'ailleurs (on peut même dire que les postulats qui caractérisent les treillis de De Morgan ne sont que les axiomes de $\mathbf{R_{fde}}$ écrits autrement).

On peut néanmoins prouver un résultat plus fort, en relation avec le treillis **4** du diagramme 2) ci-dessus, qui délivre une procédure de décision « sémantique » pour $\mathbf{R_{fde}}$:

THÉORÈME DE CARACTÉRISATION. — Une conséquence du premier degré $A \rightarrow B$ est prouvable si et seulement si elle est valide dans **4** (Dunn, 1986).

1) Si elle est prouvable, elle est valide dans **4**, puisqu'elle est valide tout court.

2) Supposons que A → B ne soit pas prouvable ; on n'a donc pas [A] ≤ [B]. Par le Théorème d'homomorphisme, il existe un homomorphisme **h** de l'Algèbre de Lindenbaum dans **4** tel que si x n'est pas inférieur ou égal à y, **h**(x) n'est pas inférieur ou égal à **h**(y). Donc dans **4**, on n'a pas **h** ([A]) ≤ h([B]), de sorte que A → B n'est pas valide dans **4**.

Le ressort de la preuve est évidemment le Théorème d'homomorphisme, qui permet de passer d'un contre-exemple à un autre. Il repose sur un autre théorème (Stone, 1937 ; non démontré ici), la propriété dite de séparation du filtre, qui énonce que :

> Dans un treillis distributif, si non-($x \leq y$), il existe un filtre premier F tel que $x \in$ F et $y \notin$ F (le filtre» sépare » les deux éléments).

Soit un treillis distributif T, et non-($x \leq y$) ; par la propriété de séparation du filtre, il existe un filtre premier F tel que $x \in$ F et $y \notin$ F. Pour tout élément t de T, on définit **h**(t) selon la place de t et de son complémentaire de De Morgan relativement à F :

(1)	$t \in$ F, $-t \notin$ F :	**h**(t) = 1
(2)	$t \in$ F, $-t \in$ F :	**h**(t) = a
(3)	$t \notin$ F, $-t \in$ F :	**h**(t) = 0
(4)	$t \notin$ F, $-t \notin$ F :	**h**(t) = b.

Ces spécifications assurent que **h**(x) ∈ {a, 1}, et **h**(y) ∈ {0, b}, comme on peut le vérifier aisément. Il resterait en toute rigueur à vérifier que **h** est bien un homomorphisme.

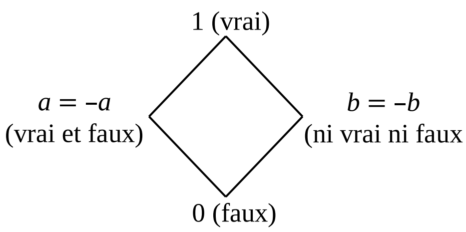
À titre d'exemple, regardons la formule $p \rightarrow (q \vee \neg q)$, qui n'est pas un théorème de **R**$_{\mathbf{fde}}$. Selon le théorème, elle est donc falsifiable dans **4**. En effet, on peut considérer la valuation **v** définie par :

v(p) = a **v**(q) = b ;

il est clair que **v** ($q \vee \neg q$) = b, et *a n'est pas* ≤ b ! En fait, dans tout treillis où la complémentation n'est pas booléenne, cette formule peut être falsifiée. Il en est évidemment de même pour la formule (A ∧ ¬A) → B.

Une sémantique à quatre valeurs de vérité ?

Il n'est pas facile de dire exactement ce qui distingue une sémantique algébrique d'une « véritable » sémantique digne de ce nom. Peut-être la suggestion suivante est-elle provisoirement acceptable : une interprétation est réellement sémantique quand elle utilise un cadre conceptuel où les notions de vérité et de fausseté sont mobilisées, d'une manière ou d'une autre. Considérons une généralisation du cadre classique, celui des tables à deux valeurs de vérité (ou de l'algèbre de Boole à deux éléments), sous la forme d'une interprétation à quatre valeurs de vérité. Le treillis **4** peut être exemplifié par la structure suivante :



Les tables des connecteurs peuvent être facilement reconstituées à partir du diagramme. La table de la négation, par exemple, est la suivante :

$\neg$	
1	0
a	a
b	b
0	1

Quant à la conjonction, on voit par exemple que $a \wedge b = 0$, $a \wedge 1 = a$, $a \wedge \neg a = a$, etc. On peut vérifier que la formule $(p \wedge \neg p) \rightarrow q$ n'est pas valide dans ce modèle, en considérant par exemple la valuation :

$\mathbf{v}(p) = a$ (vrai et faux) $\mathbf{v}(q) = 0$, ou b (faux, ou ni vrai ni faux).

On peut vérifier également que le syllogisme disjonctif : $p \wedge (\neg p \vee q) \rightarrow q$ ne l'est pas non plus.

Le problème d'une telle sémantique est d'ordre plutôt conceptuel : est-il acceptable d'introduire des valeurs de vérité comme *à la fois vrai et faux,* et *ni vrai ni faux ?* Si l'on attribue les valeurs de vérité à des entités linguistiques (énoncés, ou occurrences particulières d'énoncés), il semble relativement naturel d'admettre des énoncés ni vrais ni faux. Par exemple, un énoncé contenant un désignateur vide, *i.e.* sans référence, peut être tenu pour ni vrai ni faux (Frege, Strawson). Les choses sont beaucoup moins claires pour la valeur *à la fois vrai et faux,* bien que les « dialéthéistes » plaident en faveur de cette possibilité (Priest, 1979 ; voir ci-dessous chap. 8). Si l'on attribue la vérité et la fausseté, non plus à des entités linguistiques, mais à des contenus abstraits exprimés par les énoncés, il paraît réellement difficile d'admettre des propositions qui ne soient ni vraies, ni fausses, ou (pire) les deux à la fois. Frege, ayant admis que les langues naturelles pouvaient former des énoncés sans valeurs de vérité, tenait qu'ils n'exprimaient pas de proposition, et faisait de la possibilité d'être vrai ou faux une propriété quasi définitionnelle des propositions (position qui fut aussi celle de Russell). Si par ailleurs une proposition était à la fois vraie et fausse, sa négation le serait également, et elles auraient toutes deux la (ou les) même(s) valeur(s) de vérité. Il y a certainement là quelque chose d'étrange.

La difficulté d'attribuer à une proposition à la fois la valeur vrai et la valeur faux (ou ni l'une ni l'autre) peut justifier l'introduction d'une sémantique légèrement différente de celle vue plus haut, quoique de même nature. Ce cadre est plus approprié à une interprétation propositionnelle du calcul de la conséquence du premier degré.

La sémantique des treillis intensionnels

Revenons à l'algèbre de Lindenbaum pour la conséquence du premier degré. Puisque deux formules interdéductibles sont intersubstituables (théorème de remplacement), on peut dire que dans le contexte de $\mathbf{R_{fde}}$, elles expriment la même entité propositionnelle. On peut donc identifier, étant donné une formule A, la classe d'équivalence de A, [A], avec la proposition exprimée par A. De ce de vue, les opérations et relations définies sur les classes d'équivalence deviennent des relations et opérations sur les propositions : $\wedge$ peut être comprise comme la conjonction de deux propositions, $-$ comme une négation propositionnelle, et $\leq$ comme la relation d'implication entre propositions. Mais on a vu que dans un treillis de De Morgan, l'opération $-$ pouvait avoir des points fixes (c'est le cas du treillis **4** justement). Or, comprise comme l'opération de négation sur les propositions, cette propriété est tout à fait problématique. Comme l'écrit Michael Dunn :

> « Il est difficile d'imaginer quelque chose de plus absurde que de supposer qu'une proposition puisse être sa propre négation, et il est à peine moins absurde de penser qu'il puisse être impossible de partager exactement les propositions entre le Vrai et le Faux, une proposition étant vraie ssi sa négation est fausse, et une conjonction étant vraie ssi chaque conjoint est vrai » (Anderson et Belnap, 1975, § 18).

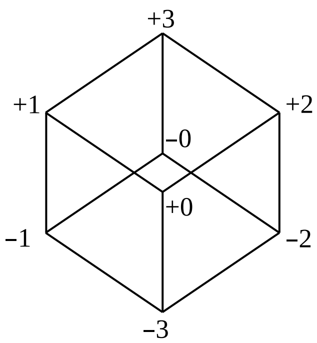
On est donc conduit à imposer, à l'opération de complémentation de De Morgan, la condition supplémentaire qu'elle n'ait pas de point fixe, *i.e.*, pour tout a d'un treillis T :

$$a \neq -a.$$

Un treillis de De Morgan pour lequel vaut cette condition est un treillis *intensionnel* (l'appellation est motivée par la remarque faite à l'instant sur ces intensions que sont les propositions). Il doit être clair que l'algèbre de Lindenbaum pour la conséquence du premier degré a une structure de treillis intensionnel, puisqu'on n'a jamais $[A] = [\neg A]$. En effet, ni $A \rightarrow \neg A$, ni $\neg A \rightarrow A$ n'y sont des théorèmes (A et $\neg$A ne sont pas interdéductibles).

On peut montrer que la condition : T n'a pas de point fixe pour la complémentation est équivalente à celle-ci : T contient un filtre F qui est à la fois consistant et complet, un « *truth filter* ». L'appellation « *filtre de vérité* » se justifie ainsi : si l'on conçoit un ensemble de propositions clos pour les opérations conjonction, disjonction, négation, avec la relation d'implication sur cet ensemble, comme un treillis intensionnel T, un filtre de vérité peut être compris comme l'ensemble des propositions vraies de T[1].

Anderson et Belnap, 1975, démontre qu'il existe un homomorphisme de tout treillis intensionnel dans un treillis intensionnel à huit élement, baptisé $\mathbf{M_0}$. $\mathbf{M_0}$ peut être représenté par le diagramme suivant :



[1] 1. L'existence d'un filtre de vérité permet de définir la notion de vérité pour toutes les *formules du premier degré, i.e.* les formules de degré zéro ainsi que toutes les fonctions de vérité de conséquences du premier degré (voir Anderson et Belnap, 1975, § 19.1)

(où -1 est le complémentaire de +1, -3 de +3, etc. ; le filtre de vérité F = {+3, +1, +2, +0}. On observera également qu'un filtre de vérité est nécessairement un filtre premier).

La considération des treillis intensionnels permet de prouver une autre forme du théorème de complétude pour $\mathbf{R_{fde}}$. Une formule A → B est prouvable si et seulement si elle est valide dans $\mathbf{M_0}$, la validité dans un treillis intensionnel étant définie comme plus haut la validité dans les treillis de De Morgan en général. La démonstration repose essentiellement sur la même propriété de séparation d'un treillis par un filtre premier, et l'existence d'un homomorphisme **h** d'un treillis intensionnel T dans $\mathbf{M_0}$, tel que pour *a*, *b*, de T, si *a* n'est pas ≤ *b*, **h**(*a*) n'est pas ≤ **h**(*b*) dans $\mathbf{M_0}$.·

Une dernière remarque pour clore ce chapitre : il est vraisemblable que l'appellation « négation de De Morgan » pour désigner ce qui peut apparaître comme une négation nouvelle, propre à la logique pertinente et distincte de la négation classique, est motivée par la sémantique des treillis de De Morgan. Mais que la négation qui figure dans *Ex Falso* ou dans le SD, soit *représentable* par l'opération unaire des treillis de De Morgan, qui ne possède pas la propriété de complémentation booléenne, ce fait *ne prouve pas*, à lui seul, que la négation de la logique pertinente soit réellement une nouvelle négation. Par la suite, j'utiliserai l'expression « négation de De Morgan » de manière purement descriptive ou inessentielle, sans que l'usage de cette expression engage à quelque décision que ce soit concernant l'identité ou la non-identité de la négation pertinente avec la négation classique.

5. LE SYSTÈME R DE L'IMPLICATION PERTINENTE

> « **§ 28.1. Pourquoi R est intéressant. E** manque d'un sens pertinent, *contingent*, de "si-alors" ; et donc, **R** promet d'avoir des applications là où ce qu'on recherche est un conditionnel dont la signification, quoique non logique, suppose la pertinence de l'antécédent pour le conséquent. Puisqu'il est certain que les locutions conditionnelles en français n'ont pas d'ordinaire de force logique, on peut s'attendre à ce que la flèche de **R**, plus souvent que celle de **E**, soit la cible appropriée pour les traductions à partir du langage ordinaire. »
>
> Anderson et Belnap, 1975.

Il s'agit à présent de combiner le calcul de l'implication pertinente et le calcul de la conséquence tautologique, pour obtenir un système qui conserve les propriétés essentielles des systèmes partiels. La réalisation de cette idée est presque exactement la suivante : on ajoute simplement au fragment implicationnel pur les axiomes pour la négation et la conséquence tautologique, à quelques ajustements près sur lesquels on va revenir.

Le langage contient cette fois tous les connecteurs propositionnels, $\neg$, $\wedge$, $\vee$, $\rightarrow$, la conjonction intensionnelle **o**, ou « fusion », et la disjonction intensionnelle $\oplus$, ou « fission », étant introduits par des définitions. À l'occasion, on peut ajouter au langage des constantes propositionnelles, munis des axiomes appropriés ; intuitivement :

t	le logiquement Vrai, conjonction de toutes les vérités logiques,
f	le logiquement Faux, disjonction de toutes les faussetés logiques,
T	le Vrai, disjonction de toutes les propositions,
F	le Faux, conjonction de toutes les propositions.

L'intérêt de l'ajout de ***t*** et de ***f*** en particulier apparaîtra plus tard, au chapitre 7.

PRÉSENTATION AXIOMATIQUE : LE SYSTÈME R

Axiomes de la conséquence (un choix possible parmi d'autres) :

R1	$A \rightarrow A$	Identité
R2	$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$	Transitivité (suffixe)
R3	$(A \rightarrow (A \rightarrow B) \rightarrow B)$	Assertion
R4	$(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$	Contraction

Conjonction :

R5 $(A \wedge B) \to A$
R6 $(A \wedge B) \to B$
R7 $((A \to B) \wedge (A \to C)) \to (A \to (B \wedge C))$

Disjonction :

R8 $A \to (A \vee B)$
R9 $B \to (A \vee B)$
R10 $((A \to C) \wedge (B \to C)) \to ((A \vee B) \to C)$

Distribution de la conjonction sur la disjonction :

R11 $A \wedge (B \vee C) \to ((A \wedge B) \vee C)$

Négation :

R12 $((A \to \neg B) \to (B \to \neg A)$ Contraposition
R13 $\neg\neg A \to A$ Double négation

Règles :

si $A \to B$, de A inférer B Modus Ponens
de A et B, inférer $(A \wedge B)$ Adjonction (ou $\wedge$Introduction)

Définitions :

D1 $A \mathbf{o} B =_{Df} \neg (A \to \neg B)$
D2 $A \oplus B =_{Df} (\neg A \to B)$

Éventuellement les axiomes pour les constantes propositionnelles :

$A \to (\boldsymbol{t} \to A)$ ***t*** Introduction
$(\boldsymbol{t} \to A) \to A$ ***t*** Élimination
$\boldsymbol{f} \leftrightarrow \neg t$
$A \to T$
$F \to A$

La présence de deux règles d'inférence est justifiée ainsi par Anderson et Belnap (ce qui vaut pour **E** vaut pour **R**, à la nécessité près) :

> « Le système **E** est destiné à couvrir deux branches de la logique formelle qui sont radicalement distinctes. La première, historiquement parlant, est concernée par les questions de pertinence et de nécessité, qui sont toutes deux à la racine des études logiques depuis les temps les plus reculés. La seconde, la logique extensionnelle, est un développement beaucoup plus récent (...). Puisque **E** couvre à la fois les deux territoires, il n'est pas surprenant que deux sortes de règles primitives soient nécessaires : la première, →Élimination, ayant à voir avec les connexions entre significations prises intensionnellement, et la seconde, &Introduction, ayant. à voir avec les connexions entre valeurs de vérité, où la pertinence n'est pas en question » (Anderson et Belnap, § 21.2.1).

Les axiomes R7 et R10 correspondent aux règles analogues pour le système des conséquences du premier degré, dont le contenu ne pouvait être exprimé par des formules dans ce cadre restreint. On peut cependant se demander s'il est réellement besoin d'une seconde règle, Adjonction, en plus du Modus Ponens. Ne pourrait-on admettre un nouvel axiome comme :

$A \to (B \to (A \wedge B))$,

tel que, par deux applications du Modus Ponens, à partir de deux formules prémisses, leur conjonction pourrait être inférée ? La réponse est *non*, sous peine de perdre la propriété de pertinence : rajouter cet axiome permet de dériver comme théorème la faute de pertinence *par excellence* : $A \to (B \to A)$.

1.	$(A \wedge B) \to A$	R5
2.	$(B \to (A \wedge B)) \to (B \to A)$	par Préfixe
3.	$A \to (B \to (A \wedge B))$	axiome proposé (?)
4.	$A \to (B \to A)$	par Transitivité

Et puisque cette dernière formule n'est pas un théorème de **R→**, l'ajout de l'axiome donnerait un système qui ne serait pas une extension conservative de **R→**, car alors une formule écrite dans le langage de **R→** ne serait démontrable que par un détour utilisant des formules avec conjonction. La formule $A \to (B \to (A \wedge B))$ n'est donc pas admissible.

L'ajout, au lieu d'un axiome, de la règle primitive Adjonction ne règle cependant pas toutes les difficultés, en ce qui concerne les déductions sous hypothèses. On peut en effet penser que des deux prémisses A, B, la conjonction $A \wedge B$ est déductible, et que de plus on a bien là une déduction *pertinente*, puisque les *deux* prémisses sont utilisées ; mais dans la perspective de l'idée sous-jacente au Théorème de la déduction, puisque B (comme A) a été *utilisée* dans la déduction de $A \wedge B$, on devrait alors avoir :

$$A \vdash B \to (A \wedge B),$$

et par une deuxième application du même théorème, la formule refusée plus haut. Autre application problématique : la déduction suivante semble admissible :

A		hypothèse
A		Répétition
	B	hypothèse
	B	Répétition
$B \to B$		→I
$A \wedge (B \to B)$		Adjonction
$A \to (A \wedge (B \to B))$		→I

d'où aussi $\vdash A \to (B \to B)$. Or ces formules violent les critères de pertinence introduits dans les chapitres précédents. Il est clair qu'il y a là un problème.

À première vue, il semble donc qu'on soit contraint de choisir entre deux directions : soit nier que de deux hypothèses A, B, leur conjonction soit dérivable, en remaniant la notion de « prémisse utilisée », ce qui reviendrait à placer une restriction sur l'emploi de la règle Adjonction (en la confinant par exemple aux théorèmes, ou en exigeant que A et B proviennent de la même formule). Soit se résigner à perdre la substance du théorème de la déduction, puisque, alors il ne serait plus vrai que chaque fois qu'on a :

$$A, B \vdash C,$$

on puisse conclure :

$$A \vdash B \to C.$$

Sinon, comme on vient de le voir, de :

$$A, B \vdash (A \wedge B), \text{ (accepté par Adjonction)}$$

on pourrait conclure :

$$A \vdash B \to (A \wedge B).$$

Il faut bien avouer que le chapitre 6 d'Anderson et Belnap, 1975, où la question est abordée, n'est pas de la plus grande clarté sur ce point[1]. Il semble que les auteurs hésitent entre les deux possibilités, choisissant clairement la première quand ils présentent la version de **R** en déduction naturelle (voir ci-dessous la restriction sur les indices des deux prémisses d'une application de la règle ∧Introduction), mais laissant le problème en suspens dans la version axiomatique. Il peut en effet paraître étrange que la conjonction $A \wedge B$ ne soit pas déductible des deux prémisses A, B, et les auteurs l'acceptent dans ce contexte axiomatique comme un pas déductif valide. Mais ils voient bien que, si on l'accepte, le prix à payer est qu'on doit abandonner la version habituelle du théorème de la déduction, comme le montre l'analyse de la « faute d'exportation », qui conduit à valider à tort la pseudo-» Loi d'exportation » (§ 22.2.2) :

$$\big((A \wedge B) \to C\big) \to \big(A \to (B \to C)\big).$$

Il est bien vrai, argumentent les auteurs, que des trois prémisses A, B, et $(A \wedge B) \to C$, C est déductible, puisqu'une application d'Adjonction donne $A \wedge B$ à partir de A et de B, d'où on obtient C par Modus Ponens. Mais « ce serait une application fautive de l'idée directrice derrière le théorème de la déduction »[2] de conclure de là que dans le contexte de A, on a $B \to C$, *i.e.* que A entraîne que B implique C. La raison invoquée en est que des prémisses sont toujours à prendre *conjointement*, de sorte que dans l'exemple d'Exportation, B n'est pas en fait une hypothèse distincte, qui pourrait être déchargée *seule et isolément.* Et c'est d'ailleurs ainsi que nous comprenons ce qu'est une collection d'axiomes en tête d'une théorie, à savoir comme ayant la valeur d'une conjonction de ces axiomes, sans qu'il y ait de suggestion de liens implicatifs entre eux. Le principe d'Exportation est donc à rejeter, précisément parce qu'il prétend inférer, d'une simple conjonction (implicite), de tels liens implicatifs. Il eût été plus simple, et plus cohérent avec le choix fait en déduction naturelle, de nier d'emblée que $A \wedge B$ soit déductible des deux prémisses A, B, quitte à justifier cette décision en décidant que la virgule n'avait, du moins pas toujours, la valeur d'une conjonction[3].

Les auteurs ne prennent pas ce parti, mais, comme c'était déjà le cas pour le fragment implicationnel, préfèrent admettre plusieurs concepts de déduction. Ils en introduisent donc un nouveau, la notion de *preuve d'une conséquence,* qui conduit à ne permettre que de décharger « en bloc » les prémisses d'une déduction, ce qui revient à donner à leur collection précisément le sens d'une conjonction. Plus précisément : on dit qu'il y a une *preuve* que des prémisses $A_1, ..., A_n$ *entraînent* une conclusion C, si et seulement s'il y a une déduction de C au sens de la version avec * (voir chap. 2), où *toutes* les

1. Voir en particulier les § 21.2.l, 22.2.2 (la discussion du principe d'Exportation), et 23.6, où est formulé un théorème de la déduction « approprié» à **E** (ou **R**), le « théorème de la conséquence».

2. Anderson et Belnap, 1975, p. 263.

3. Ce qu'Anderson et Belnap avaient par ailleurs observé dans le cadre d'une présentation en Calcul des séquents : l'antécédent d'un séquent ne peut avoir la valeur d'une conjonction, voir § 7.2.

prémisses sont marquées de *, et où, pour qu'une application d'Adjonction soit autorisée, les deux formules auxquelles s'applique la règle doivent être marquées de * (ce qui exclut qu'un des conjoints soit un théorème). Pour ce concept de déduction (disons : $\vdash$*), on a bien A, B $\vdash$* A $\wedge$ B, mais cela revient en fait à A $\wedge$ B $\vdash$ A $\wedge$ B, puisqu'une collection de prémisses est traitée comme une conjonction. Et pour ce même concept de déduction, on a une forme du Théorème de la déduction dite « Théorème de la conséquence » : s'il y a une preuve que des prémisses A_1, ..., A_n, entraînent (au sens de $\vdash$*) une conclusion C, alors $(A_1 \wedge \ldots \wedge A_n) \rightarrow C$ est un théorème de **R**. Pour prendre l'exemple le plus simple : on a bien A, B $\vdash$* A, et il est vrai que (A $\wedge$ B) $\rightarrow$ A est un théorème de **R**. Mais selon le concept « officiel » de déduction, si on admet par ailleurs que A, B $\vdash$ A $\wedge$ B, l'idée du théorème de la déduction doit être abandonnée. Et toute tentative pour le démontrer échouerait : car à partir de la déduction initiale, on ne peut inférer B $\rightarrow$ A de A pour amorcer la construction d'une preuve de B $\rightarrow$ (A $\wedge$ B) sous la seule hypothèse A, faute évidemment d'Affaiblissement.

En fait, l'option choisie en Déduction naturelle, convenablement acclimatée dans le contexte du système d'axiomes à la Hilbert, permet de démontrer une autre version du théorème de la déduction, dont le théorème de la conséquence n'est qu'un cas particulier (Kron, 1973 et 1976). Comme tout à l'heure, on exige qu'une hypothèse, pour être considérée comme réellement utilisée dans une dérivation de A $\wedge$ B, ait été utilisée pour dériver *chacun des deux conjoints* (ce qui n'était pas le cas ni de A seul, ni de B seul, dans la dérivation acceptée plus haut de A $\wedge$ B). Mais les différentes prémisses, au lieu d'être marquées de * uniformément, portent à présent des indices distincts (comme en DN). La notion de déduction pertinente étant ainsi aménagée pour tenir compte de la présence d'Adjonction, un théorème de la déduction pour **R** peut être démontré sous la forme :

$$\text{si } A_1, \ldots, A_n \vdash C, \text{ alors } \vdash B_1 \rightarrow (B_2 \ldots (B_p \rightarrow C) \ldots),$$

où les B_p, $1 \leq j \leq p$, sont, parmi les A_i, $1 \leq i \leq n$, les hypothèses *utilisées* dans la déduction initiale. La démonstration est la même que celle du théorème de la déduction pour **R→** à ceci près qu'il faut à présent tenir compte du cas où la formule traitée est obtenue par la règle Adjonction ; dans ce cas, quand on regarde dans la déduction initiale une formule D $\wedge$ E obtenue en partant de B_j par Adjonction, on peut supposer par hypothèse de récurrence $(B_j \rightarrow D)$, $(B_j \rightarrow E)$, et inférer par l'axiome R7 : $B_j \rightarrow (D \wedge E)$, qui traite donc un des cas de l'étape de récurrence[1].

Comme on le verra mieux plus tard, l'obscurité qui règne sur cette question dans *Entailment* tient pour l'essentiel au fait suivant : alors que la relation de conséquence a lieu, de manière générale, entre une collection de prémisses et une conclusion, la question de savoir comment il faut comprendre une telle collection n'est pas mise au premier plan dans cet ouvrage, dont c'est sans doute la faiblesse essentielle. S'agit-il de quelque chose d'équivalent à une conjonction, auquel cas :

$$A, B \vdash A \wedge B,$$

1. Kron, 1973, montre qu'on peut passer des deux formules $(B_j \rightarrow D)$, $(B_j \rightarrow E)$, à leur conjonction qui est l'antécédent de l'axiome R7, parce que ces deux formules portent le même indice dans la dérivation qu'on construit, ce qu'on démontre par récurrence. Les choses étant plus claires en déduction naturelle, je n'insiste pas.

mais aussi du même coup :

$$A, B \vdash A$$

sont acceptables ? Mais alors l'idée du théorème de la déduction, l'équivalence entre implication et déductibilité, doit être abandonnée, ou restreinte à la décharge simultanée des prémisses, sous forme de leur conjonction. Y a-t-il au contraire un lien plus fort entre les prémisses, celui qui assure précisément que lorsqu'on a :

$$A, B \vdash C,$$

on a aussi :

$$A \vdash B \to C,$$

i.e. on a une relation entre A et B qui fait que A entraîne que B implique C. Les choses eussent été plus claires si, au lieu de multiplier les notions de déduction, les auteurs avaient clairement vu qu'il y a plusieurs manières de comprendre ce qu'est une collection de prémisses, et distingué entre une collection extensionnellement comprise, A *et* B, et une collection intensionnellement comprise, quelque chose comme A *avec* B (sur ce point, voir le chapitre 7). L'introduction du connecteur fusion, **o**, sera un premier pas dans cette direction.

Une dernière remarque sur le système d'axiomes : le Tiers exclu, $A \vee \neg A$, est un théorème de **R** :

1.	$\neg B \to ((\neg B \to B) \to B)$	Assertion
2.	$\neg B \to (\neg B \to \neg (\neg B \to B))$	Contraposition
3.	$\neg B \to \neg (\neg B \to B)$	Contraction
4.	$(\neg B \to B) \to B$	Contraposition
5.	$(\neg B \to \neg A) \to ((\neg A \to B) \to (\neg B \to B))$	Transitivité
6.	$(\neg B \to \neg A) \to ((\neg A \to B) \to B)$	Transitivité
7.	$(A \to B) \to ((\neg A \to B) \to B)$	Contraposition
8.	$A \to (A \vee \neg A)$	axiome R8
9.	$\neg A \to (A \vee \neg A)$	axiome R9
10.	$(A \vee \neg A)$	2 Modus Ponens sur 7, 8, 9

Il s'ensuit immédiatement que toutes les formules comme $p \vee q \vee \neg p \vee r$, disjonctions d'atomes où figurent une variable propositionnelle et sa négation, et éventuellement d'autres atomes, sont prouvables dans **R** (comme dans **E**).

LE SYSTÈME DE DÉDUCTION NATURELLE DNR

Les règles avec indices souscrits pour l'implication et la négation sont les mêmes que précédemment. On a à présent pour la conjonction :

- ∧Élimination :

$$\frac{(A \land B)_{\mathbf{a}}}{A_{\mathbf{a}}} \qquad\qquad \frac{(A \land B)_{\mathbf{a}}}{B_{\mathbf{a}}}$$

- ∧Introduction (avec restriction sur les indices) :

$$\frac{\begin{matrix} A_{\mathbf{a}} \\ B_{\mathbf{a}} \end{matrix}}{(A \land B)_{\mathbf{a}}}$$

où le fait que A et B aient *les mêmes indices* relève de l'option évoquée plus haut : « Ce qui est pertinent pour une conjonction doit l'être pour chacun des conjoints » (Anderson et Belnap, 1975, § 23.1). Autrement dit, la conjonction doit dépendre d'une même prémisse, qui a été utilisée pour déduire les deux conjoints. Ce qui bloque, non seulement la déduction de $A \land B$ à partir de A, B, mais aussi la déduction de $A \land (B \lor \neg B)$ à partir de A, les théorèmes n'étant pas accompagnés d'indices.

Qu'en est-il pour la disjonction ? La disjonction est interprétée classiquement, comme (semble-t-il) la négation, et donc $A \lor B$ peut être défini comme $\neg (\neg A \land \neg B)$. Les règles d'introduction et d'élimination pour ∨ peuvent donc être introduites comme règles dérivées :

∨Introduction : de $A_{\mathbf{a}}$, inférer $(A \lor B)_{\mathbf{a}}$; de $B_{\mathbf{a}}$, inférer $(A \lor B)_{\mathbf{a}}$

1.	A_1	hyp.
2.	$(\neg A \land \neg B)_2$	hyp.
3.	$\neg A_2$	∧Elimination
4.	$(\neg A \land \neg B) \rightarrow \neg A$	→Introduction
5.	$A \rightarrow \neg(\neg A \land \neg B)$	Contraposition
6.	$\neg(\neg A \land \neg B)_1$	→Elimination

La règle ∨Élimination :

de $(A \lor B)_{\mathbf{a}}$, $(A \rightarrow C)_{\mathbf{b}}$, $(B \rightarrow C)_{\mathbf{b}}$, inférer $C_{\mathbf{a} \cup \mathbf{b}}$,

se démontre de manière analogue.

Un problème provient cependant de R11, la distributivité de la conjonction sur la disjonction. En déduction naturelle « classique », une preuve pourrait avoir l'allure suivante :

1.	$A \land (B \lor C)$	hyp.
2.	A	∧Élimination
3.	$B \lor C$	∧Élimination
4.	B	hyp.
5.	A	réitération

6.	$A \wedge B$	$\wedge$Introduction (?)
7.	$(A \wedge B) \vee C$	$\vee$Introduction
8.	$B \rightarrow ((A \wedge B) \vee C)$	$\rightarrow$Introduction (?)
9.	C	hyp.
10.	$(A \wedge B) \vee C$	$\vee$Introduction
11.	$C \rightarrow ((A \wedge B) \vee C)$	$\rightarrow$Introduction
12.	$(A \wedge B) \vee C$	$\vee$Élimination

Les pas problématiques, du point de vue de la logique pertinente, · sont marqués d'un point d'interrogation : si l'on rétablit les indices souscrits, B doit avoir un indice *nouveau* (c'est une nouvelle hypothèse, comme C ultérieurement), $\wedge$Introduction ne peut s'appliquer. Si, d'un autre côté, on tolérait que la conjonction soit déductible des deux prémisses A, B, c'est le pas 8 qui ne serait plus justifié, on l'a vu. On se retrouve à nouveau ici devant l'irritant problème de ce qu'est une collection de prémisses. Quoi qu'il en soit, prise à la lettre, la restriction sur les indices bloque la preuve. Mais comme par ailleurs la formule (axiome R11) :

$$A \wedge (B \vee C) \rightarrow ((A \wedge B) \vee C)$$

satisfait tous les critères de pertinence reconnus, la solution retenue par Anderson et Belnap, 1975, est de prendre la règle correspondante comme règle *primitive* en déduction naturelle :

Règle (primitive) de distributivité :

de $A \wedge (B \vee C)_{\mathbf{a}}$, inférer $((A \wedge B) \vee C)_{\mathbf{a}}$.

Cet état de choses peut être jugé peu satisfaisant. D'une part, la symétrie typique en Déduction naturelle des règles d'Introduction et d'Élimination est perdue. D'autre part, on peut mettre en accusation les règles du système, puisqu'elles ne permettent pas de démontrer une formule reconnue comme valide. Stephen Read, après avoir rappelé la restriction sur $\wedge$Introduction selon laquelle A et B doivent porter le même indice, écrit ainsi :

> « Anderson et Belnap imposent une restriction semblable sur les inférences concernant "ou". En conséquence, ils s'aperçoivent que leurs règles de déduction naturelle ne garantissent plus l'assertion de Distribution (...). Leur formulation en déduction naturelle est donc en désaccord avec les autres critères de validité qu'ils privilégient, selon lesquels Distribution est une conséquence du premier degré valide » (Read, 1988).

Ce genre d'objection peut conduire à une version de **R** sans la distributivité, mais je n'entrerai pas ici dans le débat pour ou contre la distributivité, qui ne concerne pas directement la notion de conséquence[1]. On notera pour l'instant juste ceci : la preuve rejetée plus haut semble suggérer qu'on ne peut la démontrer sans accepter une notion extensionnelle d'une collection de prémisses, *i.e.* un sens de A, B, où la virgule ait le sens d'une conjonction. C'est bien le cas (voir chap. 7), et ce n'est pas étonnant : après tout, $\wedge$ et $\vee$ sont des connecteurs purement extensionnels !

1. Voir sur ce point la position très balancée de Belnap dans *Lift in the Undistributed Middle* (Schroeder-Heister et Dosen, 1993). **R** sans la distributivité est décidable, alors que **R** (avec) ne l'est pas : mais on peut vivre sans la décidabilité, fait remarquer entre autres Belnap !

R CONTIENT LE CALCUL CLASSIQUE DES TAUTOLOGIES

Que les motivations qui ont mené à la construction des différents systèmes partiels restent bien intactes quand ces fragments sont combinés entre eux suppose qu'on puisse montrer que chacun des sous-systèmes est bien le fragment correspondant de **R** : que **R→** est exactement le fragment implicationnel de **R**, etc. Autrement dit, que **R** est bien une extension conservative de **R→**, et aussi une extension conservative de $\mathbf{R_{fde}}$, etc., au sens où une formule écrite dans le langage de l'un de ces sous-systèmes, mais non démontrable dans ces systèmes, ne devient pas démontrable dans **R** par un détour. On se contente ici de montrer que **R** contient le calcul propositionnel classique, et que ce dernier « occupe la place à laquelle on s'attend » (§ 24.1) ; autrement dit, le système **R** est une extension conservative du calcul classique : toutes les tautologies sont des théorèmes de **R** et, d'autre part, parmi les formules de degré zéro, *i.e.* qui ne contiennent aucune occurrence de →, seules les tautologies sont prouvables.

On montre que le calcul classique des énoncés est contenu dans **R**, en le construisant sous une forme, le système **TV** (pour « Truth values »), qui rend particulièrement aisée la preuve que toutes les tautologies classiques sont des théorèmes de ce système. Le système **TV** est une procédure de preuve, et en fait de décision, pour la propriété d'être une tautologie. Disons qu'une formule est de degré zéro *au sens strict*, si elle ne contient que des atomes (variables et négations de variables), des négations, et des disjonctions ; cela n'est pas une véritable restriction, puisque toute formule de degré zéro au sens habituel peut être transformée en formule de degré zéro en ce sens restreint, par exemple :

$$(p \supset q) \supset (r \;\&\; (q \supset p)) \text{ devient : } \neg(\neg p \lor q) \lor \neg\big(\neg r \lor \neg(\neg q \lor p)\big).$$

Les parties disjonctives d'une formule A, outre A elle-même, sont les composants sur lesquels porte une disjonction : ce peuvent être des atomes, des doubles négations, des disjonctions, ou des négations de disjonctions. Une formule contenant une partie disjonctive A sera représentée uniformément par :

(... ∨ A ∨ ...),

même si l'une ou l'autre (ou les deux) parties disjonctives sont vides.

La procédure est la suivante. Soit A une formule, dont on se demande si elle est une tautologie ; on l'inscrit comme racine d'un arbre, et on construit l'arbre en remontant :

- étape de base : A ne contient comme parties disjonctives que des atomes ; dans ce cas, A est une tautologie si et seulement si A contient une variable et sa négation ;

- étape inductive : A contient au moins une partie disjonctive qui est soit une double négation, soit une négation de disjonction ; soit B une telle partie, disons la plus à gauche. On réduit la question portant sur A à une question portant sur B, par les règles suivantes de construction d'un arbre :

1) Si B est une double négation, *i.e.* B = ¬¬C, écrire :

(... ∨ C ∨ ...) au-dessus de (... ∨ B ∨ ...) ;

2) Si B = ¬(C ∨ D), ouvrir un embranchement et écrire :

$(\ldots \vee \neg C \vee \ldots) \qquad (\ldots \vee \neg D \vee \ldots)$

au-dessus de : $(\ldots \vee B \vee \ldots)$.

Au fur et à mesure qu'on avance dans la construction de l'arbre, la complexité des formules diminue, jusqu'à ce qu'on parvienne à des disjonctions d'atomes. S'il s'agit de tautologies, on dira que ce sont des *axiomes* de **TV**. Les règles d'écriture, prises en sens inverse, constituent des règles d'inférence ; un arbre où figurent des tautologies à tous les sommets constitue donc une preuve dans **TV** de la formule racine à partir des axiomes, qui sera donc un théorème de **TV** (il est clair par ailleurs qu'il s'agit d'une procédure de décision pour la propriété d'être un théorème, et en fait, d'être une tautologie).

THÉORÈME. — Une formule est une tautologie si et seulement si elle est un théorème de **TV**.

1) Tous les théorèmes sont des tautologies (correction) :

En effet, d'une part, les axiomes sont des tautologies (des formes du Tiers exclu) ; d'autre part, les règles transmettent la propriété :

– si $(\ldots \vee C \vee \ldots)$ est tautologique $(\ldots \vee \neg\neg C \vee \ldots)$ l'est aussi, puisqu'elles ont toujours même valeur de vérité ;

– si $(\ldots \vee \neg C \vee \ldots)$ et $(\ldots \vee \neg D \vee \ldots$ sont des tautologies, deux cas sont possibles : ou pour une assignation de valeurs de vérité, $\neg C$ et $\neg D$ prennent la valeur Faux, auquel cas au moins une autre partie disjonctive de $(\ldots \vee \neg C \vee \ldots)$ et $(\ldots \vee \neg D \vee \ldots)$ prend la valeur Vrai. Dans ce cas $(\ldots \vee \neg(C \vee D) \vee \ldots)$ a également la valeur Vrai. Ou $\neg C$, et $\neg D$, sont les seules parties à prendre la valeur Vrai, auquel cas C et D sont toutes deux fausses ; alors $(C \vee D)$ a la valeur Faux, $\neg(C \vee D)$ la valeur Vrai, de sorte que $(\ldots \vee \neg(C \vee D) \vee \ldots)$ a la valeur Vrai à nouveau.

2) Toutes les tautologies sont des théorèmes (complétude) :

Soit A un non-théorème, *i.e.* il y a sur l'arbre une branche avec une formule terminale qui n'est pas un axiome. On peut falsifier toute partie disjonctive de toute formule figurant sur cette branche, en donnant la valeur Faux aux variables qui figurent sur la branche, la valeur Vrai à celles qui n'y figurent pas ; tout $\neg p$ figurant sur cette branche est falsifié, puisque si $\neg p$ y figure, p n'y figure pas (la formule terminale n'étant pas un axiome), et a donc reçu la valeur Vrai. La formule terminale prend évidemment la valeur Faux pour cette assignation. Inductivement, si C a la valeur Faux, $\neg\neg C$ aussi ; si $\neg(C \vee D)$ figure sur la branche, ou $\neg C$, ou $\neg D$, y figure aussi, et par hypothèse avec la valeur faux ; dans ce cas C (ou D) a la valeur Vrai $(C \vee D)$ également, si bien que $\neg(C \vee D)$ a la valeur Faux. Donc toute partie disjonctive de toute formule de la branche est falsifiée, la formule-racine A aussi, et A n'est pas une tautologie.

On peut prouver à partir de là que le calcul classique **TV** est inclus dans **R,** et donc que toutes les tautologies sont des théorèmes de **R.** En effet, tous les axiomes de **TV** sont des théorèmes de **R** (voir la démonstration du Tiers exclu), et les règles correspondent à des théorèmes de **R,** en fait de $\mathbf{R}_{\mathbf{fde}}$:

$\vdash (A \vee B \vee C) \rightarrow (A \vee \neg\neg B \vee C)$
$\vdash (A \vee \neg B \vee D) \wedge (A \vee \neg C \vee D) \rightarrow (A \vee \neg(B \vee C) \vee D)$.

De plus, **R** est une extension conservative de **TV,** au sens où, parmi les formules de degré zéro, seules les tautologies sont prouvables. On le montre en réinterprétant systématiquement $\rightarrow$ par le conditionnel matériel ; sous cette interprétation, tous les théorèmes de **R** sont des tautologies, de sorte que, parmi les formules de degré zéro en particulier, on ne trouve pas de théorème non tautologique. On peut conclure de ces deux points que la logique propositionnelle classique est le fragment extensionnel de **R**[1].

LA CONJONCTION INTENSIONNELLE

Dans la version axiomatique de **R** présentée plus haut, le connecteur « fusion » pour la conjonction intensionnelle a été introduit par la définition explicite :

$$A \mathbf{o} B =_{Df} \neg(A \rightarrow \neg B),$$

qui suggère une interprétation intuitive de **o** : « A et B sont compatibles », *i.e.* A n'entraîne pas la négation de B, ou A n'interdit pas que B soit le cas. Ce même connecteur peut aussi être introduit comme primitif, à l'aide des deux axiomes :

o1 $\quad A \rightarrow (B \rightarrow (A \mathbf{o} B))$
o2 $\quad A \rightarrow (B \rightarrow C) \rightarrow ((A \mathbf{o} B) \rightarrow C)$.

En fait, les deux options reviennent au même dans **R** : la définition permet de démontrer ces deux formules à titre de théorèmes, et les deux axiomes permettent de retrouver le contenu de la définition sous forme d'une équivalence (interdéductibilité). À titre d'illustration, voici une preuve de **o**1 en Déduction naturelle :

1.	A_1	hyp.
2.	B_2	hyp.
3.	$(A \rightarrow \neg B)_3$	hyp.
4.	$\neg B_{1,3}$	$\rightarrow$Élimination
5.	$\neg(A \rightarrow \neg B)_{2,1}$	Contraposition
6.	$(A \mathbf{o} B)_{2,1}$	Définition

Les indices assurent que deux applications de $\rightarrow$Introduction donnent le résultat attendu[2].

1. **R** est de même une extension conservative de $\mathbf{R_{fde}}$: si $A \rightarrow B$ est une conséquence du premier degré, qu'elle soit prouvable dépend de la prouvabilité de conséquences primitives $A_i \rightarrow B_j$. Par le théorème selon lequel, si une formule $A \rightarrow B$ est prouvable dans **R,** A et B ont une variable en commun, de tels A_i et B_j doivent avoir une variable en commun, autrement dit les $A_i \rightarrow B_j$ sont des conséquences tautologiques. Il s'ensuit que, parmi les conséquences du premier degré, seules les conséquences tautologiques sont prouvables dans **R**. On peut également considérer un système de logique positive **R+**, etc.

2. En Déduction naturelle, on peut introduire une règle **o**-Introduction : de $A_{\mathbf{a}}$ et $B_{\mathbf{b}}$, infërer $A \mathbf{o} B_{\mathbf{a} \cup \mathbf{b}}$; et une règle **o**-Élimination : d'une déduction de C à partir de $A_{\mathbf{a}}$ et $B_{\mathbf{b}}$, infërer $A \mathbf{o} B \rightarrow C$ dont l'indice ne contient ni **a** ni **b.**

Preuve de **o**2 : à partir de $A \to (B \to C)$, on a par Contraposition, $A \to (\neg C \to \neg B)$, d'où par Permutation $\neg C \to (A \to \neg B)$; par Contraposition à nouveau, et Définition :

$$(A \mathbf{o} B) \to C.$$

La réciproque de **o**2 est également démontrable, par exemple en reprenant le chemin de cette preuve à l'envers. Finalement, on a :

$$((A \mathbf{o} B) \to C) \leftrightarrow (A \to (B \to C)),$$

ou sous une forme peut-être plus parlante :

$A \mathbf{o} B \vdash C$ ssi $A \vdash B \to C$,

où $\vdash$ indique la relation de déductibilité dans l'une ou l'autre version de **R**.

Cette équivalence manifeste le lien très étroit entre **o** et $\to$, qu'on peut paraphraser intuitivement ainsi : « A *combiné* avec B implique C si et seulement si A implique que B implique C. »[1] Le lien suggéré par **o** est plus fort que celui de la simple conjonction (d'où le nom de « conjonction intensionnelle »). On le voit clairement si on contraste l'équivalence, lue de gauche à droite, avec le soi-disant « principe d'Exportation », réputé invalide :

$$\big((A \wedge B) \to C\big) \to \big(A \to (B \to C)\big).$$

L'argument contre Exportation était que la conjonction de A et de B peut bien entraîner C, *sans que cela entraîne* que A soit de conséquence sur les liens implicatifs entre B et C. L'équivalence montre au contraire que la combinaison A **o** B, si elle implique C, exprime exactement ce lien. Mais, par ailleurs, **o** n'a pas la propriété typique de la conjonction, d'impliquer chacun des conjoints. A **o** B *n'implique pas* A (ni B). Ce qu'on peut vérifier facilement, au vu de la réciproque de **o**2 : si A **o** B impliquait A, on pourrait déduire la faute de pertinence par excellence, $A \to (B \to A)$.

On peut se demander si le raisonnement usuel est sensible à cette différence. Je suggère que des exemples plausibles de conjonction intensionnelle soient fournis par les situations suivantes : à partir de lois naturelles exprimées par des conditionnels (Anderson et Belnap, 1975, les appellent des « tickets d'inférence »), combinées à une information nouvelle concernant des faits initiaux auxquels s'appliquent ces lois, nous pouvons inférer des prédictions. La loi, prise comme hypothèse, selon laquelle :

quand l'altitude augmente, l'air se refroidit,

combinée avec la nouvelle que les nuages approchent des montagnes, permet d'inférer qu'il va pleuvoir, par la condensation de la vapeur d'eau. Et la même idée est exprimée en disant que la loi par elle-même entraîne l'implication : si des nuages approchent des montagnes, il pleut. Ce qui est exactement le contenu de l'équivalence démontrée à l'instant, et tend à montrer que dans la combinaison d'une loi avec une information sur des conditions initiales, il y a bien plus que la conjonction ordinaire.

1. D'un point de vue algébrique, ce lien porte un nom, « résiduation » (voir par exemple Michael Dunn, L'algèbre de **R**, § 28.2 *in* Anderson et Belnap, 1975). Je ne développerai pas cette sémantique algébrique pour **R**, privilégiant dans la suite la sémantique relationnelle de Routley-Meyer, à des fins de discussion.

Ordinairement, la donnée des axiomes d'une théorie enveloppe sans doute au contraire la conjonction extensionnelle. Anderson et Belnap argumentent ainsi ce point :

> « Quand on dit que les axiomes de la théorie des groupes impliquent que l'élément neutre est unique, nous comprenons que leur *conjonction* l'implique. Personne ne comprendrait cette affirmation comme voulant dire que clôture, associativité et existence d'*un* élément neutre, impliquent conjointement que l'existence-d'un-inverse *entraîne* que l'élément-neutre-est-unique. Comment *pourrait-on* déduire quoi que ce soit concernant les conséquences d'un ensemble incomplet d'axiomes pour les groupes, qui *ne mentionne même* pas la notion de conséquence ? » (§ 22.2.2).

L'introduction de ce nouveau connecteur est susceptible d'éclairer au moins partiellement les problèmes relatifs à l'application de la règle Adjonction. Si l'on admet que de A, B, $A \wedge B$ est déductible, c'est qu'on interprète clairement la virgule comme une conjonction extensionnelle ; mais alors on perd le contenu du théorème de la déduction (ou $\rightarrow$Introduction) : voir le rejet d'Exportation. Si l'on refuse au contraire que de A, B, $A \wedge B$ soit en général déductible (en restreignant la règle au cas où les deux prémisses proviennent d'une même formule), c'est qu'on interprète tacitement la virgule comme **o**. Ce que confirme l'opposition entre la formule refusée au début de ce chapitre : $A \rightarrow \big(B \rightarrow (A \wedge B)\big)$ (de deux prémisses, inférer leur conjonction) et l'axiome **o**1 : $A \rightarrow \big(B \rightarrow (A \mathbf{o} B)\big)$, de deux prémisses inférer leur combinaison intensionnelle. L'introduction de **o** est donc bien un premier pas vers la clarification de ce que peut être une collection de prémisses.

LA SÉMANTIQUE RELATIONNELLE POUR R

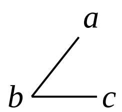
La sémantique relationnelle pour **R**, dite « de Routley-Meyer », provient d'une généralisation des idées de Kripke pour la logique modale[1]. Et de même que la sémantique des mondes possibles, elle accomplit en un sens une réduction extensionnelle de la logique pertinente. Les phénomènes d'intensionnalité sont en effet interprétés par la distribution des deux valeurs de vérités en différents indices, « *set-ups* », ou points : on parlera ainsi de la valeur de vérité d'un énoncé A au point *a*. Et la valeur de vérité de A en *a* ne dépendra pas en général seulement des valeurs de vérité de ses parties composantes en *a*, mais de leurs valeurs de vérités en d'autres points reliés à *a* par certaines relations.

Une **R**-structure de modèle (« ***R**-frame* ») est la donnée d'un ensemble non vide K de « mondes » possibles, ou plutôt de constructions (« *set-ups* ») ou points, d'un monde distingué 0 conçu comme l'analogue du monde réel, d'une relation ternaire R sur K, et d'une opération unaire * de négation, soumises à un certain nombre de conditions.

En sémantique modale, on lit d'ordinaire l'affirmation que *a*R*b* : « le monde *b* est accessible à partir de *a* », ou encore « *b* est possible relativement à *a* ». Pour la relation ternaire R, la paraphrase de R*abc* n'a rien d'immédiatement évident. Guidé par la

1. Je suis pour l'essentiel la présentation de Routley-Meyer, 1973, à quelques modifications près tirées de Dunn, 1986.

métaphore des « mondes », on peut proposer : *« a* et *b* sont compatibles du point de vue de *c* » ; en termes d'information, on dira plutôt quelque chose comme : « L'information *a* combinée à *b* est disponible en *c* » (Je reviendrai sur ces interprétations au chapitre suivant). On peut quoi qu'il en soit tenter de visualiser la relation R à l'aide de la configuration suivante :



qui représente le fait que R*abc, i.e.* que *a* est compatible avec *b* du point de vue de *c.* Cette relation permet de formuler une clause pour l'évaluation d'une formule A → B dans un modèle sur la base d'une structure de modèle :

$$v\big((A \to B), a\big) = \text{Vrai } \underline{\text{ssi}} \text{ pour tous } b, c \in K \text{ tels que R}abc,$$
$$\text{si } v(A, b) = \text{Vrai}, \qquad v(B, c) = \text{Vrai},$$

où la fonction *v* d'évaluation est une application qui à chaque paire formée d'une formule et d'un point de K, associe l'une des deux valeurs de vérité. Comme pour les sémantiques de Kripke en logique modale, les propriétés qu'on impose à la relation R permettent de distinguer différentes logiques pertinentes ; les conditions données ci-dessous caractérisent le système **R**.

Il peut être éclairant de rappeler tout d'abord les contraintes qui pèsent sur une sémantique formelle appropriée à **R** :

1) En vue d'une propriété de complétude, il faut faire en sorte que les non-théorèmes ne soient pas valides au sens de cette sémantique, et en particulier s'il s'agit de non-théorèmes de la forme A → B, que A n'entraîne pas B. Par exemple, il ne faut pas que n'importe quelle formule entraîne $q \vee \neg q$, ni non plus $q \to q$ (ni $A \to (B \vee \neg B)$, ni $A \to (B \to B)$ ne sont des théorèmes de **R**). Ce résultat est assuré en admettant la possibilité qu'une *vérité logique,* comme $q \vee \neg q$, ou $q \to q$ (qui sont des théorèmes de **R**), *ne soit pas vraie* dans tous les *set-ups,* pour toutes les valuations dans toutes les structures de modèles (qu'une vérité logique puisse ne pas être vraie quelque part justifie amplement qu'on évite de parler de « mondes » à propos des points). De manière duale, puisque $(A \wedge \neg A) \to B$ n'est pas un théorème, on admettra la possibilité qu'une contradiction soit vraie en quelque point.

2) En vue d'une propriété de correction, on veut que tout théorème soit valide. Par exemple, $q \vee \neg q$, $q \to q$ doivent être valides. Comment concilier cette exigence avec l'idée que les vérités logiques puissent ne pas être vraies en tous les points ? En définissant la validité, non pas par la vérité dans tous les mondes, mais par la vérité dans l'analogue du monde réel, 0 (ou si l'on veut, dans tous les mondes *normaux,* étant donné que le monde réel est l'un des mondes normaux : les mondes normaux sont justement les mondes où valent toutes les lois logiques). Comme l'écrivent Routley et Meyer :

> « Il est nécessaire de distinguer 0 pour la raisons suivante : la vérité logique *n'est pas la vérité* dans toutes les constructions (*set-ups*) : car la stratégie qui fait disparaître les paradoxes consiste à permettre que même les identités logiques se révèlent parfois fausses. (...) Qu'est-ce donc qu'une vérité logique ? C'est la vérité dans toutes les constructions, bien sûr, *où toutes les vérités logiques sont vraies !* » (Routley.. :Meyer, 1973 ; note : les paradoxes en question sont évidemment ceux de l'implication matérielle).

3) Il est naturel de demander que nous raisonnions conformément aux lois logiques « en vigueur dans notre monde », le monde réel, donc que si A entraîne B, A → B soit valide, *i.e.* vrai dans notre monde pour toutes les valuations, et réciproquement. Le Lemme de vérification (voir ci-dessous) montre que, sous les définitions appropriées de « entraîner » et de « valide », ce résultat est assuré (mais « valide » doit avoir une définition indépendante, dans la mesure où certaines formules valides ne sont pas des conséquences, voir le Tiers exclu par exemple).

4) Pour que $q \vee \neg q$ ne soit pas vrai dans tous les mondes, il faut admettre des points où l'assertion de q, comme celle de $\neg q$, peuvent faire défaut, *i.e.* des points incomplets. De manière duale, pour qu'une contradiction n'implique pas n'importe quelle formule, il faut admette des points où p et aussi $\neg p$ peuvent être simultanément vrais, *i.e.* des points inconsistants. On ne peut donc pas s'en tenir à la clause habituelle pour la négation : « ¬A est vrai dans un monde (pour une valuation) si et seulement si A ne l'est pas. » C'est ici qu'intervient l'opération *, qui associe à tout point a le point a^* ; ce traitement de la négation caractérise la sémantique de Routley-Meyer, et permet de formuler une clause récursive pour la négation : « ¬A est vrai dans a (pour une valuation) <u>ssi</u> A n'est pas vrai dans a^* (pour cette valuation). » Supposons un monde a où A ∧ ¬A soit vrai pour une valuation ; ¬A est donc vrai en a, donc A est faux en a^*. Mais A est également vrai en a. Si ¬A était vrai en a^*, par la clause pour la négation, A serait faux en a^{**}. En admettant que $a^{**} = a$, il s'ensuivrait que A serait faux en a, contrairement à l'hypothèse. Donc ¬A est également faux en a^* : a est inconsistant, cependant que a^* est incomplet (relativement à la négation).

Formellement, on impose à la relation R les propriétés suivantes. On définit tout d'abord :

Déf. 1 : $a \leq b = \mathrm{R}0ab$,

Déf ; 2 : $\mathrm{R}°(ab)cd = \exists x\ (\mathrm{R}abx$ et $\mathrm{R}xcd)$,

Déf. 3 : $\mathrm{R}°a(bc)d = \exists x\ (\mathrm{R}axd$ et $\mathrm{R}bcx)$.

Déf. 1. prend en compte le statut particulier de 0 : $\mathrm{R}0ab$ dit de quelque manière que a est compatible avec b *tout court.* Quant à Déf. 2, la relation R° peut être comprise comme une généralisation de la relation de compatibilité : $\mathrm{R}°(ab)cd$ dit ainsi que d est accessible à partir du triplet a, b, c, ou encore que a, b, c sont compatibles relativement à d. L'utilité de Déf. 3 apparaîtra plus loin. Les conditions imposées à R et à * sont les suivantes :

1.	$\mathrm{R}0aa$ (i.e. $a \leq a$)	Identité[1],
2.	$\mathrm{R}aaa$	Idempotence
3.	$\mathrm{R}°(ab)cd \Rightarrow \mathrm{R}°a(bc)d$	Associativité
4.	si $\mathrm{R}abc$, alors $\mathrm{R}bac$	Commutativité
5.	si $\mathrm{R}abc$, et $a' \leq a$, alors $\mathrm{R}a'bc$	Monotonie
6.	$\mathrm{R}abc \Rightarrow \mathrm{R}ac^*b^*$	Inversion
7.	$a^{**} = a$	Involution

1. Les noms des conditions sont donnés par M. Dunn en relation avec les propriétés de l'opération de monoïde de la sémantique algébrique pour **R**.

Remarque : Il s'ensuit de ces propriétés que $0^* \leq 0$, i.e. : R00*0 ; en effet on a par Idempotence : R0*0*0*, d'où R0*0**0**, d'où R0*00 (par Inversion et Involution), d'où R00*0 par Commutativité.

Soit S l'ensemble des formules du langage de **R**, et ⟨K, 0, R, *⟩ une **R**-structure ; un modèle M = ⟨K, 0, R, *, *v*⟩ est une structure où *v* est une valuation, *i.e.* une application de S × K dans {Vrai, Faux} satisfaisant les conditions suivantes :

1) *Condition héréditaire atomique :* pour tous $a, b \in$ K, et toute variable p, si $a \leq b$ et $v\,(p, a)$ = Vrai, alors $v\,(p, b)$ = Vrai.

2) Clauses pour les connecteurs :

v (A ∧ B, a) = Vrai ssi v (A, a) = Vrai et v(B, a) = Vrai ;
v (A ∨ B, a) = Vrai ssi v (A, a) = Vrai ou v(B, a) = Vrai ;
v ((A → B), a) = Vrai ssi pour tous b, c ∈ K tels que Rabc,
si v (A, b) = Vrai, v(B, c) = Vrai ;
v (¬A, a) = Vrai ssi v (A, a^*) = Faux.

On peut ajouter la clause pour le connecteur **o** :

v ((A **o** B), a) = Vrai ssi il existe $b, c \in$ K, tels que Rbca,
v (A, b) = Vrai et v (B, c) = Vrai,

mais cette clause est redondante si A **o** B est défini. Soit un modèle où A **o** B est vrai en a. Puisque B est vrai en quelque c, ¬B est faux en c^* (par la clause pour la négation). Comme on a Rbca, on a aussi Rba^*c^* (par la propriété 6), et par Commutativité, on a Ra^*bc^*. Donc la formule A → ¬B est fausse en a^*, puisqu'il existe un point b et un point c^* tels que a^* et b sont compatibles du point de vue de c^*, mais A est vrai en b, ¬B est faux en c^*. Donc sa négation, ¬ (A → ¬B), est vraie en a. Le raisonnement est analogue dans l'autre sens.

Définitions

Une formule A *est vraie en a pour une valuation v* ssi v (A, a) = Vrai, sinon fausse. On notera parfois « A est vraie en a pour une valuation » : $a \vDash v$ A.

Une formule A est *vérifiée pour une valuation v* ssi v (A, 0) = Vrai.

Une formule A est *valide dans une structure de modèle* ssi elle est vérifiée par toute valuation dans cette structure.

Une formule A est *valide* ssi elle est valide dans toutes les structures de modèles.

A *entraîne* B *pour une valuation v* ssi pour tout $a \in$ K, si $a \vDash v$ A, $a \vDash v$ B.

A *entraîne* B *dans une structure* de modèle ssi A entraîne B pour toutes les valuations sur cette structure.

A *entraîne* B (ou « ***R***-*entraîne* ») ssi A entraîne B dans toutes les structures de modèles.

La relation *sémantique* de conséquence (« ***R***-*entailment* ») ici définie, n'est évidemment pas la relation exprimée dans **R** par → ; en fait, c'est la prouvabilité dans **R** d'une formule A → B qui indique que A ***R***-*entraîne* B, de même qu'en logique classique la prouvabilité de A ⊃ B indique que A implique (classiquement) B. En outre, ni entraîner *pour* une valuation, ni entraîner *dans* une structure de modèle ne correspondent à la relation de conséquence logique ; ces. notions « caractérisent tous les arguments préservant la vérité dans des contextes spécifiques ; en prenant en compte toutes les valuations et toutes les structures de modèle, nous parvenons à la notion logique recherchée » (Routley et Meyer, 1973).

LEMME D'HÉRÉDITÉ (extension de la condition héréditaire atomique à toutes les formules) :

Si $a \leq b$ et $a \vDash v$ A, alors $b \vDash v$ A.

La condition 1) formulée plus haut sur les valuations assure que c'est vrai pour les variables. Le résultat est par ailleurs immédiat si le connecteur principal est $\wedge$ ou $\vee$. Si A $= \neg$B : supposons $a \vDash v \neg$B, pour une valuation *v*. B est donc faux en a^*, et comme $a \leq b$, $b^* \leq a^*$, par la condition 6 (Inversion). B est donc faux en b^*, sinon, en vertu de l'hypothèse de récurrence, B serait vrai en a^* ; d'où $\neg$B est vrai en *b*. Si A $=$ (B $\rightarrow$ C) : supposons $a \vDash v$ (B $\rightarrow$ C) ; soient *c* et *d* tels que R*bcd*, et supposons $c \vDash v$ B ; par hypothèse du lemme, $a \leq b$, et donc par Monotonie, R*acd ;* donc $d \vDash v$ C, puisque $a \vDash v$ (B $\rightarrow$ C) ; et donc pour tous *c, d,* tels que R*bcd,* si $c \vDash v$ B, alors $d \vDash v$ C, *i.e.* $b \vDash v$ (B $\rightarrow$ C).

LEMME DE VÉRIFICATION. — Si A entraîne B pour une valuation *v*, alors A $\rightarrow$ B est vérifié pour *v, i.e.* $0 \vDash v$ A $\rightarrow$ B. Donc, si A entraîne B, A $\rightarrow$ B est valide.

Il faut montrer que sous l'hypothèse, pour tous *a, b,* tels que R0*ab, i.e.* tels que a $\leq$ *b,* si $a \vDash v$ A, $b \vDash v$ B. Or si $a \vDash v$ A, comme A entraîne B, $a \vDash v$ B, et par le Lemme d'hérédité, $b \vDash v$ B ; donc $0 \vDash v$ A $\rightarrow$ B.

La réciproque du lemme est également vraie : supposons $0 \vDash v$ A $\rightarrow$ B ; pour tous *a, b,* tels que $a \leq b$, si $a \vDash v$ A, $b \vDash v$ B ; par Identité ($a \leq a$), on a donc : si $a \vDash v$ A, $a \vDash v$ B, *i.e.* A entraîne B.

THÉORÈME (correction). — Si A est un théorème de **R**, alors A est valide (vérifié pour toute valuation dans toute structure de modèle).

Il faut d'abord montrer que tous les axiomes de **R** sont valides. Comme les axiomes sont tous de la forme B $\rightarrow$ C, il suffit de montrer, en vertu du Lemme de vérification, que B entraîne C. Pour chaque axiome, on suppose donc que l'antécédent est vrai dans un monde *a* (pour une valuation *v* quelconque), et on montre que le conséquent est aussi vrai dans *a* (pour *v*).

La preuve est immédiate pour R1. On prouve facilement la validité de R3 (Assertion). Supposons que A est vrai dans un monde *a* : il faut montrer que (A $\rightarrow$ B) $\rightarrow$ B l'est aussi (pour une valuation quelconque). Soit *b, c,* tels que R*abc,* et (A $\rightarrow$ B) vrai en *b ;* par Commutativité, R*bac.* Par hypothèse (A $\rightarrow$ B) vrai en *b* ; A est vrai en *a,* donc B est vrai en *c.* Donc (A $\rightarrow$ B) $\rightarrow$ B est vrai en *a,* ce qui est le résultat recherché. (Noter le rôle de la condition 3 dans la preuve de la loi d'Assertion.)

Validité de l'axiome R2 (Transitivité-Suffixe) : supposons que (A $\rightarrow$ B) est vrai en *a,* il faut montrer que si R*abc* et (B $\rightarrow$ C) est vrai en *b* (A $\rightarrow$ C) est vrai en *c, i.e.* si R*cde* et A vrai en *d,* C est vrai en *e.* Supposons donc R*abc,* (B $\rightarrow$ C) vrai en *b,* A vrai en *d,* et R*cde.* Puisque R*abc,* par Commutativité, R*bac.* Il existe donc *x* tel que R*bax* et R*xde,* *i.e.* $R°(ba)de$, d'où par Associativité $R°b(ad)e$, *i.e.* il existe x tel que $Rbxe$ et $Radx$. B est donc vrai en x, et C en e, ce qu'il fallait démontrer.

Validité de R4 (Contraction) : on suppose que $A \to (A \to B)$ est vrai dans a (pour quelque valuation) : il faut montrer qu'alors $(A \to B)$ est vrai dans a, *i.e.* que pour tous b, c, tels que $Rabc$ et A est vrai dans b, B est vrai dans c. Supposons que $Rabc$ et que A est vrai dans b. Par Commutativité, $Rbac$, et comme $Rbbb$ (condition 2), on a $R°(bb)ac$, d'où $R°b(ba)c$ (condition 3). En vertu de la définition de $R°b(ba)c$, il existe x tel que $Rbax$ et $Rbxc$, d'où par commutativité à nouveau, il existe x tel que $Rabx$ et $Rxbc$. Puisque par hypothèse $A \to (A \to B)$ est vrai dans a, que A est vrai dans b, et que $Rabx$ $(A \to B)$ est vrai dans x. Comme A est vrai dans b, et $Rxbc$, B est vrai dans c, ce qu'il fallait démontrer.

Les axiomes pour la conjonction et la disjonction (R5 à R11) sont aisés à démontrer. Pour l'axiome R12 (Contraposition) : supposons que $(A \to \neg B)$ soit vrai en a, et $(B \to \neg A)$ soit faux en a (pour quelque valuation). Il existe alors b, c, tels que $Rabc$, B vrai en b, et $\neg A$ faux en c. Il s'ensuit que A est vrai en c^*, $\neg B$ faux en b^*, et Rac^*b^* (condition 6), *i.e.* que $(A \to \neg B)$ est faux en a ; contradiction. Pour l'axiome R13 : supposons que $\neg\neg A$ soit vrai en a (pour quelque valuation) ; $\neg A$ est faux en a^*, d'où A est vrai en $a^{**} = a$ (condition 7).

Il reste à montrer que les deux règles Modus Ponens et Adjonction préservent la validité. En fait, on peut montrer plus généralement : pour tout a (et pas seulement pour 0), si A est vrai en a (pour quelque valuation), et $(A \to B)$ est vrai en a, B est vrai en a, par la condition 2. Et si A est vrai en a, et B est vrai en a, il s'ensuit immédiatement que $A \wedge B$ est vrai en a.

Remarque sur le Tiers exclu : $A \vee \neg A$ est un théorème de **R,** donc doit être valide, *i.e.* vérifié dans 0 pour toute valuation. Et c'est bien le cas. En effet, ou A est vrai dans 0 (pour quelque valuation), ou $\neg A$ est vrai dans 0. Cela dépend du fait que $0^* \leq 0$, établi plus haut. Supposons en effet que A ne soit pas vrai dans 0 ; alors $0^* \vDash \neg A$, et par le Lemme d'hérédité, $0 \vDash \neg A$.

Théorème de complétude : Toutes les formules valides sont des théorèmes de **R**.

(Cette démonstration, plutôt laborieuse, ne sera pas donnée ici. Elle figure dans la version imprimée de cet ouvrage.)

6. LES INFORTUNES DE LA SÉMANTIQUE

> « La sémantique étudie les relations entre le langage et le monde. La sémantique mathématique, ou théorie des modèles, le fait en construisant des structures mathématiques (des objets abstraits de genres variés), qui, d'un côté, représentent certains aspects du monde et, de l'autre, entretiennent des relations avec les expressions du langage étudié. C'est une pratique tout à fait commune en sémantique mathématique d'identifier simplement le monde avec la structure qui le représente. Mais cette identification cache un aspect important de l'entreprise entière, si bien qu'ici nous y résistons. »
>
> Barwise et Perry, 1983, chap. 3.

Les questions qui vont faire l'objet de ce chapitre peuvent être abordées par le biais du problème posé par le syllogisme disjonctif, tel qu'il a été brièvement évoqué au chapitre 4. Ce problème peut être simplement formulé ainsi : le logicien classique tient que le SD est valide, ou encore que $\neg A \wedge (A \vee B)$ entraîne la formule B. La logique pertinente le nie. La sémantique relationnelle pour **R** construit des modèles (structures et valuations) où, au sens *d'entraîner* qu'elle définit, $\neg A \wedge (A \vee B)$ n'entraîne pas B (il y a des points a et des valuations v tels que $v\ (\neg A \wedge (A \vee B),\ a) = 1$, alors que $v(B, a) = 0$). Question : ce fait a-t-il la moindre portée en ce qui concerne l'évaluation de la validité du SD ? Peut-il être invoqué comme une preuve que, finalement, le SD n'est pas un mode d'inférence valide ?

Le syllogisme disjonctif qui est l'objet de la dispute est bien évidemment le mode d'inférence où les connecteurs qui figurent dans la prémisse $\neg A \wedge (A \vee B)$ (ou dans l'antécédent de l'implication correspondante) sont les connecteurs ordinaires, vérifonctionnels, classiques. Dans la pensée des pères fondateurs de la logique pertinente, la faute de pertinence résidait bien dans la pseudo-déduction, à partir de *non*-A (négation classique), *et* (conjonction extensionnelle), A *ou* B (disjonction extensionnelle), d'une formule B quelconque :

> « Nous nous occupons ici de la pertinence de l'antécédent pour le conséquent, quand les deux sont purement vérifonctionnels. (...)
> « En rejetant le principe du syllogisme disjonctif, nous entendons restreindre notre rejet aux cas où "ou" est pris vérifonctionnellement » (Anderson et Belnap, 1975, chap. III, § 15 et § 16.1).

Le « défi de Geach », on s'en souvient, consiste à défendre le syllogisme disjonctif en faisant remarquer que non seulement on n'a jamais présenté de contre-exemple infirmant la validité du schéma, mais qu'on ne voit pas du tout comment il serait possible d'en trouver un. Puisqu'il s'agit bien de la disjonction ordinaire, vérifonctionnelle, il est impossible que $(A \vee B)$ soit vrai, et les deux disjoints faux ; donc l'un au moins est vrai. Mais si par ailleurs la seconde prémisse $\neg A$ est vraie, la négation étant bien la négation classique, A est faux. Donc c'est B qui est vrai. Il est donc impossible que les prémisses soient vraies et la conclusion fausse.

Comme y ont insisté des auteurs comme Read et Copeland, il ne servirait à rien, de la part des défenseurs de la logique pertinente, de plaider qu'on a en fait affaire à d'autres connecteurs, distincts des connecteurs vérifonctionnels familiers, pour expliquer l'invalidité du syllogisme disjonctif (une telle stratégie serait au demeurant tout à fait à l'opposé des remarques initiales d'Anderson et Belnap, qui justement admettaient qu'avec un·» ou » intensionnel, assurant un lien de pertinence entre les disjoints, le mode d'inférence était légitime). Si l'on veut relever le « défi de Geach », c'est évidemment sur son propre terrain qu'il faut se battre : invoquer d'autres connecteurs pour montrer qu'un schéma analogue au SD, mais construit avec ces nouveaux connecteurs, n'est pas valide, est sans la moindre portée sur la question : le « bon vieux SD familier et extensionnel », est-il, oui ou non, valide ? Read écrit à ce sujet :

> « Le point de départ est l'affirmation par le logicien pertinent que le SD, ou la règle de détachement pour "$\supset$", est invalide, au sens donné par le logicien classique à "$\vee$" et à "$\supset$". Pour autant qu'il s'agit de la dispute entre pertinentiste et classiciste (*sic*), cela ne prouverait rien d'exhiber un sens·non vérifonctionnel de "ou" ou "$\supset$", selon lequel l'inférence serait invalide. Ce qu'il faut au logicien pertinent, c'est justifier sa thèse que le logicien classique se trompe au sujet du bon vieux SD familier et extensionnel. Le défi de Geach, qu'il ne le peut pas, est toujours là » (Read, 1988).

Les mêmes remarques valent bien sûr pour la négation. Copeland, dans un article qui justement conteste la portée de la sémantique de Routley-Meyer sur la question de la valeur du SD, fait également remarquer que :

> « L'invalidité [*the failure*] de la formule $\sim A \wedge (A \vee B) \rightarrow B$ dans une sémantique qui donne un sens non classique à l'un quelconque des connecteurs $\wedge$, $\vee$, $\sim$, n'a aucune portée sur la question clef de savoir si une conséquence pertinente a lieu entre les formules $\sim A \wedge (A \vee B)$ de la logique propositionnelle, et la formule B - c'est-à-dire n'a aucune portée sur la question de savoir si le syllogisme disjonctif commet ou non une faute de pertinence. Routley et Meyer ont donc la charge de démontrer que leur nouveau traitement de $\sim$ dans leur sémantique préserve son sens classique » (Copeland, 1979).

C'est bien là le point. La sémantique des **R**-structures de modèle n'est de conséquence sur la question de la validité du SD que si l'on peut montrer qu'elle respecte le sens des connecteurs qui y figurent, et en particulier qu'elle donne à la négation son sens classique. Ce n'est nullement évident. D'une part - raison purement factuelle -, parce que la littérature pertinente elle-même parle volontiers de « négation pertinente », ou de « négation de De Morgan » à propos du connecteur interprété à l'aide de l'opération *, pour l'opposer à la négation classique[1]. D'autre part, et surtout, parce que

1. Dans ce contexte, la négation de De Morgan est fréquemment notée « $\sim$ », la négation classique, booléenne, « $\neg$ ». L'appellation « de De Morgan » provient entre autres du fait qu'elle satisfait les lois de De Morgan (mais évidemment ni Ex Falso, ni le SD) ; dans le cadre de la sémantique algébrique pour **R**, celle des semi-groupes réticulés, la négation est interprétée par une opération ayant les mêmes propriétés que l'opération vue au chapitre 4 sur les treillis de De Morgan. Voir sur ce point Anderson et Belnap, 1975, § 28.2, l'Algèbre de **R**.

la question des propriétés que doit avoir une négation — acceptons l'idée d'une pluralité de négations —, pour être réputée classique, cette question est loin d'être claire.

En fait, au-delà du problème particulier de la négation, une question plus générale se pose. Dans leur présentation de la sémantique, Routley et Meyer définissent la notion fondamentale : A *est vrai en a pour une valuation v,* par : v (A, a) = Vrai (ou 1, si les valeurs possibles de v sont prises dans {0, 1})[1]. Dans quelle mesure cette notion est-elle susceptible de donner aux connecteurs leur signification attendue, en termes de conditions de vérité ? Et d'abord dans quelle mesure sommes-nous justifiés à interpréter cette notion comme un équivalent formel du concept préformel de vérité ? D'autres interprétations semblent possibles, et rien ne dit, après tout, que l'interprétation *aléthique* soit cohérente avec tous les aspects de la sémantique relationnelle. Là encore, la réponse est loin d'être évidente, parce que la question de savoir comment décider si une interprétation est admissible, ou dans quelle mesure la construction sémantique la contraint, n'est pas facile à résoudre.

INTERPRÉTER LA SÉMANTIQUE FORMELLE

Dans une critique argumentée de la sémantique de Routley-Meyer, « What is a Semantics for Classical Negation ? », Copeland insistait sur le fait qu'une sémantique purement mathématique (formelle) ne peut remplir, à elle seule, la fonction qu'on peut attendre d'une véritable sémantique : assigner une signification aux symboles d'un langage[2]. Pour que les clauses récursives qui en sont le cœur, par exemple, expliquent véritablement la signification des logiques, il faut que la propriété sémantique, dont les clauses donnent les conditions de sa possession par une expression complexe, en termes de sa possession par les expressions composantes, soit de quelque manière antérieurement comprise. Typiquement, étant donné une compréhension préalable du prédicat de vérité, les clauses récursives permettent d'articuler le sens attendu des connecteurs logiques. Il est· donc nécessaire, pour que ce travail soit accompli, que les propriétés ou relations fondamentales du formalisme de la sémantique pure soient connectées avec des concepts sémantiques familiers liés à nos usages linguistiques :

> « Pour qu'une signification soit assignée par une sémantique récursive aux constantes logiques, il est nécessaire que le formalisme mathématique de la sémantique soit compris en termes d'un concept sémantique antérieurement compris - en ce cas, la vérité. (...)
> « La conversion d'une sémantique pure en une sémantique appliquée consiste donc en l'interprétation de la récursion mathématique comme établissant des conditions nécessaires et suffisantes pour la possession, par des énoncés de la forme concernée, d'une propriété Φ

1. On peut aussi (voir Dunn, 1986) caractériser un modèle par la donnée d'une relation ⊨ entre points et formules, si bien que les valuations peuvent être identifiées aux fonctions caractéristiques de ces relations. Bien que par facilité on lise volontiers « a ⊨ A », a rend vrai A, ou A est vrai dans a, le même problème d'interprétation de cette relation se pose. Dans ce qui suit, je parlerai aussi bien de ⊨ que de valuation pour la relation sémantique fondamentale.

2. Dummett oppose de la même façon la construction de structures abstraites qui peuvent servir d'outils d'étude d'une logique, à la construction d'une véritable sémantique (Dummett, 1977, § 7.3 en particulier). Je ne discuterai pas ici de la question des deux buts qu'on peut assigner à une sémantique : définir un prédicat de vérité (version Tarski), ou utiliser un prédicat de vérité pour fixer le sens des opérateurs logiques (version, disons, Davidson) : Sur cette opposition, voir Etchemendy, 1988, et Rivenc, 1998.

antérieurement comprise. Φ doit être reliée à l'activité linguistique d'une manière telle que la saisie des conditions sous lesquelles les énoncés de la forme concernée possèdent cette propriété soit suffisante pour une compréhension des éléments cibles des énoncés » (Copeland, 1986).

Il importe· ici d'insister sur deux aspects de la sémantique des **R**-structures de modèle :

1) La définition inductive de la relation fondamentale ⊨ entre points et formules, qui définit la donnée d'un modèle, n'est pas amorcée par une clause initiale concernant les formules atomiques, puisque nous sommes dans le cadre d'un calcul propositionnel. Autrement dit, nous n'avons rien de tel qu'une clause initiale du type :

$a \vDash p$ si et seulement si ...

La situation est différente, donc, de ce qui se passe avec les sémantiques pour le calcul des prédicats « à la Tarski », où la définition de la notion sémantique fondamentale, en l'exemple celle de satisfaction, commence par une *définition explicite* pour les formules atomiques, grâce à des clauses initiales comme (si les clauses initiales sont formulées en relation avec des prédicats déterminés) :

l'assignation s satisfait la formule « Rxy » si et seulement si $s(x)$ R $s(y)$,

où R est l'interprétation du symbole de prédicat» R » (ce n'est pas la lettre des formulations de Tarski, mais c'en est à peu près l'esprit). C'est ici, à la base de la construction, que la connexion est faite entre le concept de satisfaction à définir formellement, et la notion sémantique usuelle (du moins pour la pensée axiomatique) de satisfaction[1]. Rien de tel ici.

2) Comme les éléments de K sont des objets abstraits, *set-ups,* points, ou indices absolument quelconques dont nous ne savons rien, la donnée d'un modèle, c'est-à-dire d'un ensemble arbitraire de couples (a, p), ne nous dit rien à elle seule sur la nature de cette relation. C'est pourquoi, pour que la donnée d'un modèle puisse devenir une véritable sémantique, il faut bien, suivant la formule de Copeland, « connecter le formalisme mathématique avec des concepts sémantiques préexistants reliés au langage considéré ». Il doit être clair que, par elle-même, une clause récursive comme celle pour ∧ :

$a \vDash (\mathrm{A} \wedge \mathrm{B})$ <u>ssi</u> $a \vDash \mathrm{A}$ et $a \vDash \mathrm{B}$,

est incapable d'attribuer une signification au connecteur. Pour qu'une telle clause puisse être dite donner les conditions de vérité de la conjonction, il faut que nous ayons de bonnes raisons de penser que la condition $a \vDash \mathrm{A}$ et $a \vDash \mathrm{B}$ nous dit que a rend vrai A et rend vrai B. À cette fin, il faut que la sémantique formelle soit elle-même *interprétée*, en termes d'une propriété (ou d'une relation) déjà connue, telle que la compréhension des conditions sous lesquelles la formule composée possède cette propriété (ou entretient

1. Au moment d'introduire la notion de satisfaction d'un énoncé ouvert, Tarski écrit : « Essayons de clarifier par quelques exemples la signification usuelle de cette notion dans son usage linguistique habituel » (§ 3 du « Wahrheitsbegriff», *in* Woodger, 1956). La sémantique informationnelle d'Urquhart énonce une clause initiale pour les variables propositionnelles, mais qui, par elle-même, ne connecte pas la relation formelle avec une notion préthéorique de vérité ou de « détermination» (voir ci-dessous la sémantique des «stocks d'information » .

cette relation) suffise à assurer une compréhension de la signification du symbole « cible » de la formule. Il faut donc qu'une interprétation soit proposée pour la relation fondamentale $\vDash$ qui figure dans la clause. En termes de valuation, cela veut dire qu'il est nécessaire que :

> « la valuation centrale de la sémantique soit interprétée en termes d'une telle propriété, et ceci présuppose une explication de la nature des objets qui constituent le codomaine de cette fonction, des indices qui figurent dans le domaine de la fonction, des opérations et relations sur ces indices, etc. » (Copeland, 1986).

Copeland poursuit en faisant remarquer que les clauses récursives font peser une double contrainte sur toute interprétation possible de la relation fondamentale. D'une part une contrainte générale : il faut que la propriété (ou la relation) sémantique qui interprète l'appareil des clauses récursives soit telle que la possession de cette propriété par une formule composée puisse être formulée en termes de la possession de cette propriété (relation) par ses sous-formules immédiates : qu'elle soit *itérative.* Toute propriété ou relation dont on pourrait penser qu'elle est susceptible d'interpréter la notion fondamentale n'est pas itérative : pour anticiper sur la suite, la propriété *appartenir à un ensemble de croyances* n'est pas telle qu'on puisse déterminer les conditions de possession de cette propriété pour un énoncé composé, en termes de la possession de cette même propriété par ses expressions composantes. La vérité, au contraire, est la propriété itérative par excellence : les conditions pour la possession de cette propriété par une formule complexe peuvent être construites en termes de la possession de cette même propriété par les expressions composantes.

La deuxième contrainte est plus spécifique. Quelles interprétations de la relation fondamentale permettent de penser que les clauses récursives donnent aux connecteurs leur sens attendu, notamment leur sens classique ? Il est en effet permis de penser que les clauses, interprétées d'une manière ou d'une autre, donnent quelque signification aux connecteurs, même si dans certains cas cette signification est loin d'être claire (si dans certains cas extrêmes, aucune signification n'est donnée, et donc aucun connecteur véritablement défini, une telle interprétation est exclue). Une interprétation, qui donne aux connecteurs leur sens attendu, peut être dite correcte (*faithfull*) relativement à ces connecteurs.

> « Quand la transition d'une sémantique pure à une sémantique appliquée est accomplie dans un contexte où nous sommes réellement en possession d'une compréhension de la signification des connecteurs cibles de la sémantique (ou du moins de certains d'entre eux), il y a une contrainte évidente mais cruciale sur les propriétés qui sont des valeurs appropriées de Φ [*i.e.* les propriétés aptes à interpréter sémantiquement la propriété formelle utilisée dans les clauses récursives] : la propriété dans les termes de laquelle la récursion est interprétée doit être telle que la sémantique réussit effectivement à assigner leur signification attendue aux connecteurs cibles » (*ibid.*).

D'où la question : quelles interprétations de la relation $\vDash$ (étant entendu que la sémantique purement formelle laisse ouverte la porte à différentes interprétations envisageables) sont susceptibles de donner aux connecteurs leur sens attendu, en particulier leur sens classique[1] ?

1. L'École australienne favorise l'interprétation en termes de mondes incomplets et/ou inconsistants, et les théories paraconsistantes. Belnap comme Dunn ont toujours tenu à distinguer l'interprétation

L'INTERPRÉTATION PSYCHOLOGIQUE

Si l'on reprend la clause pour la conjonction rappelée plus haut, il est clair que l'interprétation de la relation fondamentale en termes de *rendre vrai* est correcte pour la signification usuelle de la conjonction. Elle articule cette signification en termes de conditions de vérité, et peut être facilement paraphrasée ainsi : *a* rend vraie la conjonction A et B ssi *a* rend vrais les deux composants. Notons que cette interprétation aléthique est solidaire d'une interprétation *ontologique* des points de l'ensemble K : ce sont des mondes, ou du moins des situations, qui sont susceptibles de rendre vrais des énoncés. Car l'interprétation de la relation est évidemment solidaire d'une certaine conception des éléments de son domaine, autrement dit des « cas » sur lesquels on quantifie, lorsque l'on dit, pour caractériser la validité : « Dans tous les cas où les prémisses sont vraies, la conclusion l'est aussi. » Mais la question est de savoir si cette interprétation est la seule possible, et même, plus gravement, si elle est correcte (au sens spécifié par Copeland), pour l'ensemble des connecteurs.

Dans « The Semantics of Entailment », les auteurs présentaient de manière plutôt libérale (ou faut-il dire désinvolte ?) différentes interprétations concevables de leur formalisme :

> « Le mot "mondes" ne nous retiendra pas, et nous utilisons la terminologie de Routley consistant à parler de *set-ups* plutôt que de mondes, indiquant par là que ce dont nous traitons n'est pas nécessairement réalisé ni même réalisable en quelque sens ordinaire que ce soit. (...) Pour faire le lien avec des intuitions qui rendent les logiques pertinentes philosophiquement intéressantes, à côté des points [*set-ups*] prosaïques qui peuvent compter aussi comme mondes, nous voudrions inclure ce qui peut être approximativement décrit comme des ensembles cohérents de croyances (quoique non nécessairement consistants), ce qui peut être présumé sur la base d'une certaine combinaison de lois et de rapports d'observation (...), ce qui serait le cas si tout ce qui devrait être fait l'était. Nous laissons de côté ici la question de savoir si la notion d'un point [*set-up*] appartient ultimement à l'ontologie, l'épistémologie, ou même simplement juste à la psychologie » (Routley et Meyer, 1973).

En reprenant de manière plus systématique les suggestions de Routley et Meyer, je me propose de considérer successivement : 1) l'interprétation psychologique en termes de croyances ; 2) l'interprétation aléthique en termes de mondes ou de situations ; 3) l'interprétation épistémologique en termes de théories. La conclusion négative sera qu'aucune de ces possibilités ne donne une interprétation cohérente de la sémantique relationnelle. Dans un second temps, on envisagera d'autres tentatives de solution.

Prenons donc au sérieux l'idée que les éléments de K sont des ensembles de croyances, et comme tels possiblement inconsistants (un individu peut bien avoir des croyances contradictoires) et très naturellement incomplets. Sous une telle interprétation du domaine, la relation $a \vDash A$ peut être comprise comme : A figure au nombre des croyances de *a*, ou tout simplement : $A \in a$. L'idée était que la signification attendue des connecteurs contraignait les interprétations possibles de la relation fondamentale qui figure dans les clauses récursives. Sous une telle interprétation psychologique, le sens

---

épistémologique (être tenu pour vrai) et l'interpétation « sémantique » ou aléthique « être vrai ». On peut aimer les logiques pertinentes sans être paraconsistant !

classique des connecteurs est-il préservé ? Il n'est pas difficile, à la suite de Copeland, de montrer que non. Considérons la clause pour la disjonction :

$a \vDash (A \vee B)$ <u>ssi</u> $a \vDash A$ ou $a \vDash B$.

Sous cette interprétation, elle veut dire (en identifiant les points et les sujets des croyances) que *a* croit (A ∨ B) si et seulement si *a* croit A, ou *a* croit B. Pour la disjonction classique, cette clause est inacceptable, dans les deux sens du « si et seulement si » : un individu peut croire (A ∨ B) sans être capable d'avoir la moindre opinion, ni sur A, ni sur B (c'est sans doute le cas le plus fréquent de croyance en une disjonction, alors que nous ignorons quel disjoint est réalisé) ; et un individu peut croire A sans croire (A ∨ B), n'ayant peut-être jamais pensé à B, ou ignorant sa signification, etc. (Copeland, 1986). À supposer que cette clause ainsi interprétée assigne quelque sens que ce soit au connecteur, ce n'est pas le sens classique. Et il n'est même pas certain que la clause parvienne à assigner quelque sens que ce soit au connecteur, parce que la propriété d'appartenir à un ensemble de croyances n'est pas *itérative* au sens proposé par Copeland. Les conditions sous lesquelles un énoncé composé est un élément d'un ensemble de croyances ne peuvent pas en général être formulées en termes de l'appartenance de ses énoncés composants à ce même ensemble de croyances. L'interprétation psychologique de la relation fondamentale semble donc exclue pour une sémantique qui se veut, en un sens élargi, *véri-conditionnelle,* même si les conditions de vérité d'un énoncé en un point dépendent éventuellement de valeurs de vérité à d'autres points, comme c'est le cas avec les connecteurs intensionnels.

Faut-il donc revenir à l'interprétation ontologique des *set-ups,* et à l'interprétation aléthique de la relation fondamentale ⊨, comme étant l'interprétation imposée par la contrainte de préserver le sens attendu, classique, des connecteurs ?

L'INTERPRÉTATION ALÉTHIQUE

La clause pour la négation (notée ici ~) peut être formulée ainsi[1] :

$a \vDash {\sim}A$ <u>ssi</u> $a^* \nvDash A$,

où * est une opération unaire censée associer à tout point *a* de K son « complementary set-up *a** ». L'explication heuristique proposée est la suivante : ce qui est fortement affirmé en un point *a,* sous la forme de la présence d'une formule A, est seulement faiblement affirmé en *a*,* sous la forme de l'omission de sa négation. Une interprétation psychologique semble ici tout à fait exclue : on voit mal à quoi correspondrait l'idée qu'un individu possède, à côté de l'ensemble de croyances où figure l'idée que la Terre se meut, un ensemble « complémentaire » d'où est absente la croyance que la Terre est immobile. L'interprétation aléthique semble s'imposer. Question : quelle signification est alors attribuée par cette clause au connecteur dit de négation ?

1. La notation ~ est souvent justifiée par l'idée que, après tout, on a bien affaire à une. nouvelle négation, la négation de De Morgan.

La conséquence de la clause (c'est à cette fin qu'elle a été inventée) est qu'il est possible qu'il existe un point a où à la fois $a \vDash {\sim}A$ et $a \vDash A$, un point réputé « inconsistant ». Il suffit pour ce faire qu'on ait $a^* \nvDash A$, car alors $a \vDash {\sim}A$; et qu'on ait aussi $a^* \nvDash {\sim}A$, car alors, selon la clause, $a^{**} \vDash A$, et moyennant le postulat sur * selon lequel $a^{**} = a$, $a \vDash A$. En d'autres termes, un point est « inconsistant » si et seulement si son compagnon par l'opération * est incomplet pour la négation.

L'interprétation aléthique de la clause est-elle correcte (*faithfull*) pour la négation classique ? Il semble à première vue que non : dans un monde où deux formules A et ~A sont vraies ensemble, ~A ne peut être dit la négation classique, c'est-à-dire le contradictoire, de A, si l'on admet que le contradictoire d'un énoncé et cet énoncé ont des valeurs de vérités opposées, et que c'est cette notion qui caractérise la négation classique. Pour autant que cette clause caractérise une négation, ce n'est donc pas la négation classique (c'est peut-être une autre forme de négation, la négation dite « de De Morgan »). Copeland écrit à ce propos :

> « *Prima facie,* il semblerait que la négation ne reçoit pas une interprétation classique chez Routley et Meyer. Car que ~ soit non classique, cela ne s'ensuit-il pas simplement du fait que $A \vee B$ et ~A peuvent être vrais à la fois dans un modèle, et B faux dans ce même modèle ? En effet, quelle meilleure preuve pourrait-il exister de la nature non classique de ~ que le fait que A et ~A puissent être vrais dans le même modèle (un trait qui semble montrer que ~ n'a en fait que très peu de rapports avec la *négation,* dont la non-contradiction est une propriété fondamentale) » (Copeland, 1979).

Sans entrer dans la question de savoir si cette nouvelle négation est, plutôt que la négation classique, la vraie négation, « la négation fondamentale du langage ordinaire », comme certains pertinentistes l'affirment, un point troublant doit être noté, dont on peut se demander s'il ne recouvre pas tout simplement un sophisme. On déclare qu'il y a des éléments de K *inconsistants,* parce que A et ~A y sont vrais ensemble. Et on tire argument de ce fait pour prouver l'adéquation du formalisme, en faisant remarquer qu'il modélise précisément soit la notion abstraite de situation inconsistante, soit plus fortement l'existence dans notre monde de situations inconsistantes. Mais il n'y a aucune raison de décréter qu'un monde est inconsistant, simplement parce qu'y figurent une formule et la même formule préfixée d'un certain connecteur ~, à moins qu'on ait, par ailleurs, *des raisons indépendantes* de penser que ce connecteur exprime la notion de contradictoire. À défaut, tout au plus peut-on dire qu'un tel monde est, en un sens purement technique, *inconsistant relativement à ce connecteur.* Copeland remarque à ce sujet :

> « Dans cette sémantique, des formules de la forme A&~A se trouvent être seulement des simili-contradictions [*counterfeit contradictions*], puisque A et ~A peuvent être vraies ensemble dans un monde possible. Des gens qui insistent sur le fait qu'il ne peut y avoir de contradiction vraie seront d'accord pour dire que A&~A peut être vrai. De ce point de vue, la négation de De Morgan ne peut d'aucune façon être invoquée en faveur des prétentions des philosophes dialectiques [*i.e.* qui l'invoquent à l'appui de la thèse qu'il existe des mondes inconsistants, c'est-à-dire contradictoires] » (Copeland, 1989).

La première conclusion qui semble s'imposer est que l'interprétation aléthique n'est pas correcte pour la négation classique. La seconde conclusion est que la négation n'étant pas la négation classique (du moins sous cette interprétation), le fait que A et ~A puissent être vrais en même temps n'est d'aucune conséquence sur la question de la validité du syllogisme disjonctif, non plus au reste que sur celle de *Ex Falso* $(A \wedge \sim A) \rightarrow B$, puisque la question posée était celle de leur validité *pour le sens classique* des connecteurs de négation, conjonction et disjonction.

Mais ne peut-on argumenter autrement ? Conclure, du fait que A et ~A peuvent être vrais ensemble dans un monde possible, que ~A n'exprime pas le contradictoire de A, c'est simplement prendre trop au sérieux la métaphore des « mondes ». Bien sûr, un point où un énoncé et son contradictoire sont vrais ensemble ne saurait être un *monde possible,* si l'on entend par là la totalité cohérente et complète d'une réalité concevable. Mais parlons en termes d'états de choses partiels, ou de situations, qui d'évidence ne déterminent pas la totalité des aspects possibles d'un monde. Les situations, pour ainsi dire par définition, sont incomplètes : la situation où j'écris à mon bureau ne détermine ni qu'il pleut en Australie, ni qu'il ne pleut pas en Australie[1]. Pourquoi ne pas admettre du même coup qu'il puisse y avoir des situations incohérentes ?

Greg Restall, par exemple, a tenté de mettre la sémantique des situations de Barwise et Perry au service d'une défense de l'opération *, en avançant l'idée que la notion de situation offrait une interprétation naturelle de la clause pour la négation (Restall, 1999 et 2000). Il y aurait beaucoup à dire sur le détail de ses arguments, mais un seul point me retiendra ici. Les auteurs de la sémantique des situations prennent grand soin de distinguer les situations réelles, qui sont (peu importe au juste de quelle manière) des parties de la réalité, et les objets et structures ensemblistes de la théorie, comme les situations abstraites, qui peuvent être actuelles (*i.e.* correspondre à une situation réelle), factuelles (*i.e.* qui classifient une situation en en épinglant un aspect), ou non factuelles. Dans le modèle mathématique (les « structures de situations »), il y a bien, pour diverses raisons, des états de choses ou des cours d'événements incohérents. Mais seuls des états ou des événements cohérents peuvent classifier des situation réelles, c'est-à-dire les décrire partiellement, *a fortiori* y correspondre. Une contrainte fondamentale de la théorie, liée au principe de non-contradiction, est qu'il n'y a pas d'états de choses *actuels* incohérents[2]. Invoquer la théorie des situations pour plaider en faveur d'une *interprétation* de la sémantique relationnelle qui repose sur l'idée de situations inconsistantes, c'est confondre le modèle mathématique et ses objets abstraits, et la réalité qu'il s'agit de modéliser. Je ne crois donc pas qu'on puisse utiliser la sémantique des situations pour défendre l'idée que, dans la réalité, il y a des situations inconsistantes, qui pourraient servir de points d'appui à une interprétation naturelle des structures de modèle abstraites.

On peut du coup être tenté de revenir à l'interprétation psychologique envisagée plus haut : car alors la conséquence de la clause pour la négation, suivant laquelle on peut trouver a tel qu'à la fois $a \vDash \sim A$ et $a \vDash A$, ne semble plus poser de problème particulier. Elle n'impose pas un sens non classique à la négation, car le fait qu'un individu possède des croyances contradictoires n'a rien que de très ordinaire. Le point a peut donc représenter un ensemble inconsistant de croyances, au sens habituel du terme, pour

1. Nombreux sont les Australiens parmi les supporters des logiques pertinentes.

2. Le texte le plus net à ce sujet se trouve au chapitre V, p. 96, de Barwise et Perry, 1983 : « La première [distinction] est celle entre situations réelles et situations abstraites. L'incohérence ne peut apparaître que parmi les dernières. » Les lignes suivantes précisent que seules les situations abstraites cohérentes peuvent être actuelles.

autant que la négation est supposée avoir son sens classique. On est là, si l'on veut, dans une situation symétrique de celle rencontrée tout à l'heure avec la disjonction. Pour la disjonction, c'était l'interprétation aléthique qui était correcte pour la disjonction classique, l'interprétation psychologique était inacceptable, et si elle donnait une signification quelconque au connecteur, ce n'était pas la signification attendue. Pour la négation, c'est l'interprétation aléthique qui est inacceptable pour la négation classique, c'est l'interprétation psychologique qui est, disons prudemment, *compatible*, avec la signification classique.

Compatible, oui : mais est-ce à dire qu'elle est « correcte » au sens défini plus haut, *i.e.* qu'elle réussit à assigner aux connecteurs leur signification attendue ? Il se peut qu'une interprétation tolère une certaine lecture des connecteurs, sans du tout la donner à proprement parler à travers les clauses récursives. Et on ne voit pas bien en quoi la clause pour ~, interprétée de manière psychologique ; est susceptible d'assigner un sens quelconque au connecteur. La question évoquée plus haut de l'existence d'un ensemble « complémentaire » de croyances mise à part, une clause qui entraîne que A et ~A peuvent appartenir à un même ensemble de croyances est bien loin de donner quelques conditions de vérité que ce soit à ~A en termes de conditions de vérité de A, et ne contraint pas ~A à être le contradictoire de A. La remarque de Copeland à ce sujet me paraît pleine de bon sens :

> « En dépit de l'absence d'interprétation de *, il est possible de commenter la clause (~) selon les perspectives des deux interprétations examinées de la sémantique (...). Selon la première façon de comprendre le formalisme [*i.e.* l'interprétation psychologique] - et en laissant de côté la question de l'existence -, $I(A,j) = T$ et $I(\sim A,j) = T$ ensemble, affirment simplement que A et ~A sont en même temps éléments de *j*, où *j* représente les croyances d'un individu, ou une théorie donnée ; etc. Avoir A et ~A simultanément acceptés, ou simultanément affirmés dans quelque théorie, ce n'est visiblement pas la même chose qu'avoir A et ~A simultanément vrais » (Copeland, 1986).

Il est donc difficile de penser qu'ainsi comprise, la clause pour la négation *assigne* une signification au connecteur ~, même si son interprétation classique est compatible avec cette clause. C'est plutôt parce que ce dernier est déjà classiquement compris, mais de manière tout à fait indépendante de la clause, qu'une croyance où figurent A et ~A est réputée inconsistante. En fait, loin que la clause accomplisse son travail sémantique, c'est la signification préexistante de la négation qui donne sens et portée à la clause, ou du moins à ses conséquences concernant la possibilité de points (croyances) inconsistants. On retombe ici sur la difficulté initiale évoquée pour commencer : la relation *figurer dans un ensemble de croyances* n'est pas itérative. Mais si l'on revient subrepticement à l'interprétation des éléments de K comme mondes, et à l'interprétation·aléthique de la relation sémantique fondamentale, car elle seule semble donner des conditions de vérité, on garde cependant du *flirt* avec l'interprétation psychologique l'idée que c'est bien la négation classique qu'on a en main (c'est elle seule qui justifie qu'on parle de croyances inconsistantes lorsqu'elles contiennent A et ~A). C'est ce va-et-vient - ou ce double jeu, je le crains -, entre les diverses interprétations sémantiques, qui donne quelque apparence de vraisemblance aux explications de Routley-Meyer sur la négation.

Une variante de l'argument des situations inconsistantes invoque celles-ci pour soutenir l'idée que la clause pour la négation, malgré les apparences, garde à celle-ci son sens classique. La négation de De Morgan, dite aussi « normale », n'est pas une nouvelle négation, mais simplement l'extension et la *généralisation* de la négation classique à de nouvelles situations, qualifiées d'inconsistantes précisément parce que la négation - réputée classique - d'un énoncé y vaut (*holds*) en même temps que l'énoncé nié. Voici l'une des expressions les plus claires de cette ligné de défense :

> « La classe des situations pour lesquelles (...) la négation est exclusive et exhaustive [*i.e.* où on a : ~A vaut dans une situation *a* ssi A ne vaut pas dans *a*] est appelée la classe classiquement contrôlée par cette négation. La variation disponible dans les classes contrôlées fournit une autre dimension selon laquelle les négations et les négations normales peuvent varier, une dimension déterminable[1] de plus dans la négation. La négation classique est simplement le cas spécial où *toutes* les situations sont classiquement contrôlées par la négation, alors que même avec une négation entièrement normale [*i.e.*, pour laquelle il existe au moins quelque situation où la négation est exclusive et exhaustive] seules certaines situations doivent être classiquement contrôlées. Ainsi une explication de la négation comme entièrement normale n'abandonne pas, mais simplement généralise l'explication classique » (Routley et Meyer, 1982, § 9).

Ce n'est donc pas la négation qui a changé, mais la classe des situations où elle s'applique. Il demeure encore une fois quelque chose d'obscur dans cet argument : si A et son « contradictoire » ont même valeur de vérité dans une situation, d'où tient-on qu'il s'agit bien du *contradictoire* de A ? La construction syntaxique à l'aide d'un connecteur ~ n'y suffit évidemment pas. La seule réponse possible est que A et ~A expriment des états de choses intrinsèquement contradictoires, mais supposés réalisés ensemble. Mais « supposés réalisés » en quel sens ? Ontologiquement, à titre de composants de situations, ou épistémologiquement, à titre d'états de choses simplement visés par une théorie ? Comment faut-il comprendre, par exemple, les « situations inférentielles » évoquées dans le texte suivant :

> « Le problème était d'expliquer ce qui n'allait pas avec le principe traditionnel du Syllogisme disjonctif, $A \mathbin{\&} (\neg A \vee B) \rightarrow B$. Après tout, le principe fut largement accepté comme correct durant quasiment les mêmes deux mille ans où les fautes de pertinence étaient réputées reconnues. Comme les critiques l'ont maintes fois fait remarquer, à chacune de ses redécouvertes, le principe n'est pas, apparemment, un cas de non-pertinence (*irrelevance*), car le conséquent est une sous-formule de l'antécédent. On ferait mieux de raconter l'histoire ainsi : ce qui ne va pas avec le Syllogisme disjonctif concerne la mise en défaut de l'assomption d'arrière-fond de consistance (locale). Il est tacitement présupposé que A exclut toujours sa négation ~A, quelque chose qui d'évidence ne tient pas dans des situations inconsistantes. En termes de faute, le Syllogisme disjonctif est une faute de consistance. Il est universellement invalide, parce qu'il ne vaut pas dans des situations inférentielles inconsistantes, et il est largement accepté comme valide, parce qu'on se restreint implicitement à des situations consistantes » (Norman et Sylvan, 1989).

Pour continuer la revue des interprétations concevables, examinons donc l'hypothèse où les « situations inférentielles » évoquées seraient en fait des théories[2].

---

1. Sur cette notion de déterminable, voir ci-dessous le chapitre 8.

2. L'extrait de *Relevant Logics and their Rivals* cité plus haut illustrait de même l'idée de situations inconsistantes ainsi : « C'est en particulier un fait qu'il existe des théories déductives, et des situations où l'on raisonne déductivement, qui sont inconsistantes, *i.e.* où A et ~A valent également » (Routley et Meyer, 1982, § 9). On y retrouve la même équivoque concernant les « situations » : états de choses réalisés ou théories ?

L'INTERPRÉTATION ÉPISTÉMOLOGIQUE

Il est possible que ce soit cette interprétation, en termes de théories (les **R**-théories), que les inventeurs de la sémantique relationnelle aient eu en vue de manière préférentielle, comme le suggère au reste l'usage de ces théories dans la preuve de complétude :

> « Si on examine le résidu syntaxique des sémantiques modales normales (...), les structures modales de Kripke peuvent être vues naïvement comme ayant des théories consistantes et complètes comme éléments. Nos structures de modèles, à l'examen de leur résidu syntaxique, se révèlent aussi avoir des théories comme éléments. Mais (...) l'attention à la pertinence exige que nous considérions des théories anormales du point de vue classique. La classe des théories que nous considérerons sera, étant donné cette contrainte, aussi normale que possible ; ses éléments seront ce que nous appelons ici des *théories intensionnelles premières* » (Routley et Meyer, 1973).

Quelques brefs rappels du chapitre précédent : une (**R**-)théorie T est *première* ssi, si $A \vee B \in T$, alors ou $A \in T$, ou $B \in T$. Une théorie est *régulière* ssi elle contient tous les théorèmes de **R**. Toute théorie n'est pas nécessairement régulière, contrairement à ce qui se passe en logique classique, où pour toute formule A, (A → théorème) est un théorème. De plus, il doit y avoir dans les structures de modèles des théories premières non régulières, dans la mesure où toute théorie première régulière est complète (relativement à la négation). En effet, si T est régulière, elle contient $A \vee \neg A$, qui est un théorème de **R**, et si en outre elle est première, alors elle contient A ou contient $\neg A$, *i.e.* est complète. Par ailleurs, les théories qui peuvent figurer comme éléments des structures de modèle doivent être premières, pour que la clause pour la disjonction soit satisfaite dans ·les deux sens du·» si et seulement si ». Si $A \in T$, $A \vee B \in T$, puisque $A \to (A \vee B)$ est un théorème de **R** ; mais dans l'autre sens, la clause impose que les théories soient premières.

Sous cette interprétation où $\vDash$ signifie à nouveau l'appartenance ; l'argument de Copeland contre le caractère inacceptable de la clause pour la disjonction tombe. L'opération * a même une interprétation relativement plausible : a étant une théorie première, a^* est l'ensemble des formules dont la négation n'appartient pas à a, $a^* = \{A\ ;\ \neg A \notin a\}$. Une théorie en ce sens étant une idéalisation de théories réelles, il n'y a peut-être pas de raison de refuser l'existence de théories « complémentaires » pour chaque **R**-théorie. Mais la relation sémantique fondamentale étant à présent comprise en termes d'appartenance à des ensembles de formules, on ne voit pas bien en quoi l'existence de situations déductives, ou « inférentielles » (théories) inconsistantes, serait de nature à invalider le syllogisme disjonctif, en dépit d'une affirmation comme celle-ci :

> « Il suffit pour falsifier le Syllogisme disjonctif qu'il y ait des théories et des situations déductives non triviales [au sens où tout énoncé n'y est pas prouvable] mais inconsistantes » (Routley et Meyer, 1982, § 2.11)[1].

1. Il ne s'agit pas seulement de la possibilité de construire des théories formelles inconsistantes mais non triviales. L'idée est plutôt qu'il y a *de fait* des théories historiquement attestées, philosophiques, mathématiques, ou physiques, inconsistantes, mais qu'on doit tenir pour non triviales, ou telles que l'usage d'une logique paraconsitante permet de montrer qu'elles ne sont pas triviales. On trouve par exemple ce genre de justification plutôt circulaire dans Priest, Routley et Norman, 1989.

Car pourquoi une telle théorie peut-elle contenir simultanément A et ~A comme thèses, sans que tout énoncé B y soit prouvable, et sans cesser pour autant d'être une *théorie*, c'est-à-dire close pour une certaine relation de conséquence ? Uniquement parce qu'elle admet comme logique sous-jacente une logique qui a rejeté $(A \wedge \sim A) \rightarrow B$ (et le SD) comme n'étant pas un théorème : autrement dit, parce qu'il s'agit précisément d'une **R**-théorie. Mais invoquer l'existence, ou du moins la possibilité, de telles théories, à titre d'argument contre la validité logique de ces modes d'inférences est manifestement circulaire. Cela revient à peu près à dire : voyez, le syllogisme disjonctif est invalide, puisque certaines théories refusent de l'admettre !

Le seul moyen de parer cette objection serait d'affirmer qu'une théorie *doit* parfois admettre deux énoncés A et ~A comme thèses, tout simplement parce qu'elle doit les tenir pour vrais tous les deux, et que le motif de son refus de déduire de leur conjonction un certain B est que B est tout simplement faux (exemple proposé : A est l'affirmation que A est faux, supposée vraie, ~A l'affirmation contradictoire également vraie, et B l'affirmation fausse qu'il n'y a pas de glace au Pôle Sud). Mais là on passe de la construction de théories inconsistantes non triviales (paraconsistance) à la réalité de situations « dialectiques », où des énoncés paradoxaux, c'est-à-dire d'authentiques contradictions, sont supposés vrais, et qui plus est vrais dans notre monde. Le *dialéthéisme* est la position qui soutient précisément cette thèse[1] ; naturellement, une telle position suppose qu'on ait de bonnes raisons de penser qu'il y a réellement des contradictions vraies dans notre monde, le monde réel. Cette analyse de la négation sera brièvement examinée au chapitre 8.

Conclusion d'étape : invoquer l'existence ou la possibilité de théories inconsistantes non triviales est sans pertinence pour la question de la validité logique *objective* du SD — si c'est cette question qu'on a en vue[2]. Invoquer des situations où une formule et sa négation sont vraies ensemble, soit disqualifie cette négation comme négation classique, soit disqualifie les modèles comme modèles *réalistes* de notre monde, à moins de soutenir la thèse dialectique que la réalité est contradictoire. Ce qui, comme le fait remarquer Read, est de toute façon une vision des choses parfaitement étrangère aux motivations conceptuelles initiales du projet des logiques pertinentes.

---

1. Priest et Routley expliquent ainsi cette terminologie : « Le mot "paraconsistant" (voulant dire "par-delà le consistant") a été introduit par Miro Quesada (en 1976) pour les théories inconsistantes mais non triviales. Pour exprimer l'idée que certaines théories inconsistantes sont *vraies* (...) nous avons été contraints d'introduire un nouveau terme. Une contradiction vraie est une sorte de Janus à double face, à la fois vraie et fausse. Le mot "dialétheia" ("vérité double") semblait tout à fait approprié pour exprimer cette idée. Ainsi le dialéthéisme est la thèse qu'il y a des *dialetheiae,* des contradictions vraies » (Priest, Routley et Norman, 1989, Introduction).

2. Mais peut-être n'était-ce pas la question que Routley et Meyer avaient en vue en construisant leur sémantique. Car, comme ils s'en sont expliqués ailleurs, l'univers des mondes possibles est une « peinture pour aider l'imagination », et ne saurait être une justification (Meyer et Sylvan, 2003, p. 380-381, *in* Brady, 2003). Il semble que Restall, au contraire, soit disposé à voir dans la notion de situation une véritable justification de la logique pertinente.

LA SÉMANTIQUE DES STOCKS D'INFORMATION

Alasdair Urquhart a proposé en 1972 une autre sémantique pour les logiques pertinentes, la sémantique dite « opérationnelle », ou encore « des semi-treillis d'information »). Ce qui distingue dès l'abord cette sémantique, c'est le soin mis à la motiver par une interprétation naturelle. Il s'agit bien, en effet, de :

> « développer une analyse sémantique naturelle des logiques pertinentes, qui refléterait plus ou moins directement les intuitions préformelles sous-jacentes à ces systèmes » (Anderson et Belnap, 1992, § 47, rédigé par Urquhart).

En particulier, la nature des « points » est explicitée en termes de stocks d'information (*piece of information*) ; la structure de semi-treillis, pour une opération $\cup$ sur les stocks d'information, est fondée sur l'analyse de ce concept : on peut réunir deux états d'information X et Y pour former leur union, $X \cup Y$. L'état d'absence de toute information est le 0 du semi-treillis. Autrement dit, l'ensemble des stocks d'information est clos pour une opération $\cup$, dont on peut admettre qu'elle est idempotente, commutative et associative, et que 0 en est le plus petit élément, *i.e.* $X \cup 0 = X$, pour tout X (noter qu'au contraire l'union de deux mondes possibles n'est pas nécessairement un monde possible). De cette structure informationnelle, on abstrait donc celle d'un semi-treillis pour l'opération $\cup$, avec 0 comme élément zéro. Sur la base de ces interprétations, la sémantique *pour le fragment implicationnel pur* de **R** peut être naturellement construite.

Un stock d'information peut être pensé comme un ensemble d'énoncés de base atomiques, dénués de complexité logique (exemples : égalités ou inégalités numériques, protocoles d'expériences en physique). Cette notion d'information doit être clairement distinguée des notions, respectivement : de situation de justification (*evidential situation*), ensemble de propositions établies au cours d'une enquête, qui doit évidemment satisfaire une exigence de consistance ; et d'autre part de monde possible ou de description d'un tel monde, qui doit satisfaire une exigence de complétude. Un stock d'information peut n'être ni consistant, ni complet. Urquhart formule clairement l'idée développée plus haut, que la nature des points d'une structure de modèle, ainsi que l'interprétation de la relation entre points et formules, sont de conséquence sur la portée des clauses récursives : à telle interprétation conviennent telles clauses, sauf à modifier le sens des connecteurs. Il écrit ainsi :

> « La distinction entre les trois concepts [information, justification, monde possible] émerge de deux manières ; d'abord, dans les structures mathématiques qu'il est naturel de considérer par abstraction à partir de ces idées ; ensuite, dans les conditions de vérité qui ont lieu pour les énoncés complexes relativement à ces structures, étant donné les conditions de vérité pour les énoncés de base » (*ibid.*).

Un stock d'information X est dit « déterminer », c'est-à-dire entraîner, un énoncé atomique *p*, si l'on est en droit de conclure *p* sur la base des énoncés contenus dans X. On serait en droit de s'inquiéter d'une apparente circularité : on veut définir la relation A *entraîne* B, et pour ce faire on postule une relation de conséquence entre ensembles d'énoncés et énoncés. Mais c'est ici qu'intervient le fait qu'un stock d'information ne contient que des énoncés de base, donc dépourvus de structure logique propositionnelle. Les relations primitives de conséquences sont donc *logic free*, en ce sens qu'elles ne reposent pas sur la structure logique liée à la présence de connecteurs. Urquhart donne comme exemples : des relations de conséquence fondées sur la synonymie de prédicats (célibataire = non marié), ou fondées sur des propriétés de relations (transitivité de « plus grand que »), supposées connues dans le contexte d'un argument.

Ces considérations justifient la sémantique formelle suivante. Sur la base d'un semi-treillis S, un modèle M est une valuation *v* qui, à chaque paramètre *p*, assigne un sous-ensemble *v*(*p*) de S (intuitivement : l'ensemble des stocks d'information qui déterminent *p*). On peut donc définir récursivement la relation $X \vDash_{(M)} p$ en commençant par la clause initiale :

$X \vDash_{(M)} p$ <u>ssi</u> $X \in v(p)$;

à lire donc : dans le modèle M, X détermine *p* <u>ssi</u> X est dans l'ensemble des stocks d'information qui déterminent *p* (la paraphrase est évidemment circulaire)[1].

Venons-en d'emblée au point crucial, l'analyse de l'implication $A \rightarrow B$: que peut vouloir dire qu'un stock d'information X détermine que A implique B ? Manifestement, si sur la base de X nous savons que $A \rightarrow B$, et si nous rajoutons à X une nouvelle information Y qui justifie A, la combinaison de X et de Y permet d'inférer B ; et réciproquement. D'où la clause pour $\rightarrow$:

$$X \vDash_{(M)} A \rightarrow B \ \underline{\text{ssi}} \text{ pour tout } Y, \text{ si } Y \vDash_{(M)} A, \text{ alors } X \cup Y \vDash_{(M)} B.$$

(Pour faire le lien avec la présentation de Routley-Meyer, on peut dire que la relation ternaire R reçoit ici l'interprétation suivante : RXYZ si et seulement si $Z = X \cup Y$, *i.e.* la combinaison des deux stocks d'information X et Y). Cette clause est fidèle à l'intuition sémantique selon laquelle le connecteur d'implication exprime la déductibilité : le rajout de l'hypothèse A permet de déduire B de $A \rightarrow B$; et réciproquement, si une nouvelle information a permis de déduire B de A, on peut conclure $A \rightarrow B$ indépendamment de cette information. Si on réinterprète les X, Y, etc., comme les indices souscrits figurant dans les preuves en Déduction naturelle pour **R→**, on retrouve exactement les règles sur les indices : une application du Modus Ponens sur deux formules d'indices X et Y donne une formule B d'indice $X \cup Y$, et si B a été déduite à partir de A d'indice Y, on peut inférer $A \rightarrow B$ d'indice X. Urquhart écrit à ce propos :

> « Nous avons, pour ainsi dire, donné une vie propre aux indices souscrits, au-delà du rôle de marqueurs qu'ils ont dans les formulations en termes de sous-preuves. Ce que nous avons ainsi souligné, c'est le caractère naturel, et la plausibilité philosophique, de l'exigence » (Anderson et Belnap, 1992, § 47 .1).

Ces deux clauses suffisent pour le fragment implicationnel de **R**. On peut définir à présent la notion de formule valide : A est *valide* <u>ssi</u> $0 \vDash_{(M)} A$ dans tous les modèles M.

Naturellement, il faut prendre l'idée d'une information qui détermine un énoncé en un sens fort : l'information doit être strictement pertinente. Ce point motive l'exclusion d'une clause qui dirait de façon générale que si $X \vDash A$, $X \cup Y \vDash A$. Techniquement, cette exclusion est cruciale pour invalider une formule comme $A \rightarrow (B \rightarrow A)$, comme on peut le voir aisément. Supposons un instant que nous admettions cette clause.

1. Ce n'est pas exactement la définition de « modèle» par Urquhart, que j'ai modifiée légèrement pour maintenir la distinction entre structure de modèle et modèle. Urquhart parle de *c*-modèle pour une paire formée de S et d'une valuation.

Supposons que $X \vDash A$; selon la clause, pour tout Y, $X \cup Y \vDash A$, et donc si $Y \vDash B$, $X \vDash (B \rightarrow A)$, d'où $0 \vDash A \rightarrow (B \rightarrow A)$. En revanche, on peut montrer que les axiomes de **R→** sont valides au sens de cette définition.

Contraction : $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$. Il faut montrer que pour tout modèle M :

$$0 \vDash_{(M)} (A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B).$$

Soient M et Y tels que $Y \vDash_{(M)} (A \rightarrow (A \rightarrow B))$, *i.e.* pour tout Z tel que $Z \vDash_{(M)} A$, $Y \cup Z \vDash_{(M)} (A \rightarrow B)$. Supposons $Z \vDash_{(M)} A$; on a donc $(Y \cup Z) \cup Z \vDash_{(M)} B$, et donc $Y \cup Z \vDash_{(M)} B$, par Associativité et Idempotence de l'opération $\cup$. Comme $Z \vDash_{(M)} A$, et $Y \cup Z \vDash_{(M)} B$, par la clause pour l'implication, on a $Y \vDash_{(M)} (A \rightarrow B)$, *i.e.* $Y \cup 0 \vDash_{(M)} (A \rightarrow B)$, ce qu'il fallait démontrer.

Permutation : montrer que $0 \vDash_{(M)} (A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$.

Soit M et Y tels que $Y \vDash_{(M)} (A \rightarrow (B \rightarrow C))$; il faut montrer que $Y \cup 0 = Y \vDash_{(M)} (B \rightarrow (A \rightarrow C))$.

Supposons que pour un Z, $Z \vDash_{(M)} B$; il faut donc montrer $Y \cup Z \vDash_{(M)} (A \rightarrow C)$.

Par hypothèse, pour tout X, si $X \vDash_{(M)} A$, $Y \cup X \vDash_{(M)} (B \rightarrow C)$ Puisque $Z \vDash_{(M)} B$, on a donc $Y \cup X \cup Z = Y \cup Z \cup X \vDash_{(M)} C$, donc finalement : si $X \vDash_{(M)} A$, $Y \cup Z \vDash_{(M)} (A \rightarrow C)$, et $Y \vDash_{(M)} (B \rightarrow (A \rightarrow C))$.

Transitivité (Préfixe) : montrer que $0 \vDash_{(M)} (A \rightarrow B) \rightarrow ((C \rightarrow A) \rightarrow (C \rightarrow B))$.

Soit Y tel que $Y \vDash_{(M)} (A \rightarrow B)$; supposons que $X \vDash_{(M)} (C \rightarrow A)$; il faut montrer que $Y \cup X \vDash_{(M)} (C \rightarrow B)$, donc que, soit Z tel que $Z \vDash_{(M)} C$, $Y \cup X \cup Z \vDash_{(M)} B$. Puisque $X \vDash_{(M)} (C \rightarrow A)$, $X \cup Z \vDash_{(M)} A$. Mais puisque $Y \vDash_{(M)} (A \rightarrow B)$, $Y \cup X \cup Z \vDash_{(M)} B$.

La démonstration de la validité de $A \rightarrow A$ est évidente, ainsi que le fait que la seule règle d'inférence (le Modus Ponens) préserve la validité. On peut donc conclure que tous les théorèmes de **R→** valides au sens de la sémantique opérationnelle[1].

Les problèmes de cette sémantique surgissent avec le traitement de la négation, dès que l'on tente de l'étendre à **R** tout entier. Urquhart note qu'évidemment, une clause de type standard « classique » :

$$X \vDash_{(M)} \neg A \text{ \underline{ssi} ce n'est pas le cas que } X \vDash_{(M)} A$$

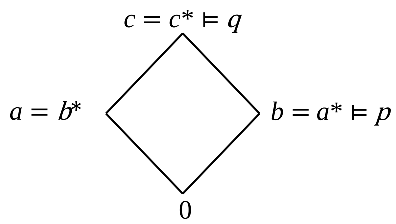
ne convient pas, car elle aurait pour effet de valider les formules « paradoxales » dont on ne veut pas. Mais il ajoute avec raison - ce qui est plus important ici - qu'une telle clause serait incohérente avec l'interprétation du système sémantique en termes de stocks d'information :

> « En d'autres termes, en adoptant cette règle, nous excluons automatiquement des stocks d'information inconsistants - on peut ajouter que c'est précisément sur cette exclusion que l'argument standard pour la validité du syllogisme disjonctif est basé. Cependant il n'y a rien, semble-t-il, dans la notion d'un stock d'information qui exige que nous restreignions notre attention aux seuls stocks consistants. De plus, nous devons admettre aussi des stocks incomplets d'information (...) sous peine de rendre valide la formule $p \rightarrow (q \vee \neg q)$ » (Urquhart, 1972).

1. La réciproque (complétude) est également vraie, mais ne sera pas démontrée ici.

Techniquement, la manière la plus simple de représenter dans le formalisme la possibilité de stocks d'information inconsistants est évidemment d'ajouter à la structure de semi-treillis S une opération * pour laquelle S est clos, telle que $0^* = 0$ (l'état zéro d'information n'est donc pas inconsistant), et telle que $X^{**} = X$. On voit facilement que la dernière condition rend valide par exemple, $A \to \neg\neg A$, et $\neg\neg A \to A$ (Double Négation Élimination).

L'un des problèmes soulevés par ce choix est qu'un des axiomes de **R** pour la négation, Contraposition, cesse d'être valide : on peut trouver un modèle où $0 \nvDash (A \to \neg B) \to (B \to \neg A)$. Soit le semi-treillis $S = \{0, a, b, c\}$ à quatre éléments, représenté par le diagramme suivant, et la valuation partiellement indiquée par la mention d'une variable propositionnelle déterminée par un élément (l'absence d'une variable signale sa fausseté) :



c vérifie $(q \to \neg p)$, puisque $c = c^*$ vérifie $\neg p$; cependant il y a un élément, nommément b, tel que b détermine p, et $c \cup b = c$ ne vérifie pas $\neg q$. Donc il est faux dans ce modèle que pour tout Y, si $Y \vDash (q \to \neg p)$, $0 \cup Y = Y \vDash (p \to \neg q)$. Donc Contraposition n'est pas valide. Il en est du reste de même pour Réduction : $(p \to \neg p) \to \neg p$ est falsifiée par ce modèle (considérer l'élément $a = b^*$). Il peut paraître surprenant que Contraposition soit ici invalidée, alors qu'elle est valide selon la sémantique de Routley-Meyer. Mais sa démonstration y reposait en fait sur le Postulat Inversion, selon lequel : si $Rxyz$, alors Rxz^*y^* (voir chap. 5). On pourrait penser à ajouter cette condition (malgré son caractère « purement *ad hoc* », *dixit* Urquhart) à la sémantique opérationnelle, en réinterprétant la relation ternaire $Rxyz$, voir plus haut. Malheureusement, on obtenir alors le résultat suivant :

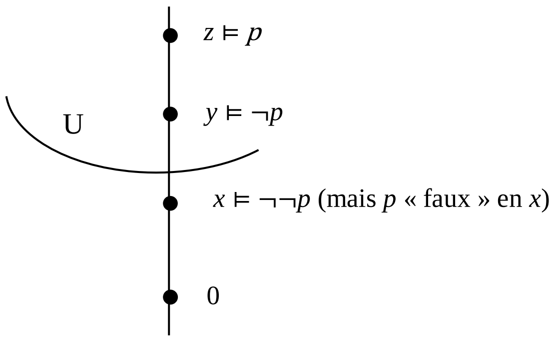
comme :	$0 \cup X = X \cup 0 = X$, on a :	$RX0X$,
	d'où :	RXX^*0^* par Inversion,
	d'où :	RXX^*0 car $0^* = 0$,
	et donc :	$X \cup X^* = 0$.

Autrement, dit $X = X^* = 0$, et il n'y a qu'un seul stock d'information dans les modèles, l'information nulle, de sorte qu'on obtient la logique classique. Cette tentative pour regagner Contraposition donc être abandonnée.

Urquhart propose donc de considérer une autre extension de sémantique des semi-treillis. On distingue un sous-ensemble arbitraire U de S, et la clause pour la négation est formulée ainsi (relativement à un modèle M) :

$X \vDash \neg A$ <u>*ssi*</u> pour tout Y, si $Y \vDash A$, alors $X \cup Y \in U$.

U peut être compris comme l'ensemble des stocks d'information contradictoires, ceux qui « déterminent » toujours une constante du faux, disons ***f***, de sorte que cela revient à définir $\neg$A comme A $\rightarrow$ ***f***. On peut donc s'attendre à ce que les formules inacceptables du point de vue intuitionniste deviennent invalides pour cette négation. Et c'est bien le cas. Les formules acceptables du point de vue intuitionniste sont bien valides : (A $\rightarrow \neg$A) $\rightarrow \neg$A est valide (sous la définition de la négation comme A $\rightarrow$ ***f***, ce n'est qu'un cas particulier de Contraction[1]). Il en est de même pour Contraposition intuitionniste, qui, sous la même définition, est un cas particulier de Permutation. Mais c'est à présent $\neg\neg$A $\rightarrow$ A, A $\vee \neg$A, et la loi de De Morgan $\neg$(A $\wedge$ B) $\rightarrow \neg$A $\vee \neg$B, qui sont invalidées. Contre-exemple pour $\neg\neg$A $\rightarrow$ A :



Pour montrer que 0 $\nvDash \neg$ (A $\wedge$ B) $\rightarrow \neg$A $\vee \neg$B, il suffit de trouver un diagramme avec, d'une part, un X tel que X $\vDash \neg$(A $\wedge$ B), et donc tel que pour tout Y qui détermine à la fois A et B, X $\cup$ Y $\in$ U. Mais il faut que X $\nvDash \neg$A $\vee \neg$B, donc ne vérifie ni l'un ni l'autre ; donc, d'autre part, il faut un point Z qui détermine A (et pas B) tel que X $\cup$ Z $\notin$ U, et un point T qui détermine B (et pas A) tel que X $\cup$ T $\notin$ U. Une telle construction est évidemment possible.

La conclusion de ces diverses tentatives était plutôt pessimiste :

> « En fait, il ne semble pas y avoir de manière acceptable de modifier les conditions de l'un ou l'autre type de négation de manière à valider tous les axiomes de **E** ou de **R**. (...) Il ne semble pas y avoir de possibilité d'incorporer les traits des deux types (sauf à ajouter des stipulations *ad hoc*). Si le cadre général présenté ici est accepté, on est conduit à la conclusion (semble-t-il à l'auteur) que toute tentative de combiner les deux dans une logique pertinente est incohérente. Ce point de vue n'était certainement pas celui des premiers articles sur les logiques pertinentes (par exemple, Ackermann, Anderson et Belnap). Cependant, aucune justification sémantique explicite n'était donnée dans ces articles pour les axiomes postulés pour la négation. Si le point de vue du présent article est correct, aucune justification ne peut être donnée » (Urquhart, 1972 ; voir aussi Urquhart, 1992 : « Il n'y a pas de construction appropriée pour la négation dans ce cadre sémantique »).

Si donc un choix s'impose entre deux variétés de négation, on peut se demander si l'interprétation de la sémantique en termes d'information motive telle négation plutôt que telle autre. Urquhart fait remarquer que la validité de la loi du Tiers exclu est tout à

1. Supposons que X F (A~ $\neg$A), il faut montrer que X F $\neg$A ; supposons que pour un certain Y, Y FA On a donc X V Y F $\neg$A, donc, en vertu de la clause pour la négation, pour tout Z tel que ZFA, XvYvZe U. Comme YFA, XvYvY=XvYe U, et donc X F $\neg$A).

fait inacceptable, c'est-à-dire l'idée que, dans tout modèle, $0 \vDash A \vee \neg A$. Vu la clause pour la disjonction, cela voudrait dire que, sur la base d'un état zéro d'information, nous serions quand même toujours capable, pour toute proposition, de conclure soit elle, soit sa négation. L'interprétation en termes d'information justifierait donc la négation minimale, mais non la négation « semi-classique », semi-classique au sens où elle justifierait le Tiers exclu, mais pas, évidemment, EFQ :

> « Ainsi nous ne devrions pas nous attendre à ce que la loi du Tiers exclu soit valide dans une sémantique mettant en œuvre des stocks d'information - bien que nous devrions nous attendre à ce qu'elle soit valide dans une sémantique utilisant seulement des mondes possibles. C'est le second type de négation, la négation minimale, qui semble bien avoir un fondement intuitif correct »[1] (*ibid.*).

La sémantique d'Urquhart est certainement une sémantique interprétée, fondée sur des considérations informelles plausibles. Mais ce n'est pas une sémantique pour la négation de **R**, et - plus grave -, Urquhart suggère qu'il est impossible de trouver une telle sémantique de manière cohérente. Quelles conclusions peut-on tirer de cet état de choses ?

LA SOLUTION DE STEPHEN READ

Dans *Relevant Logic,* Read proposait un diagnostic concernant l'origine des problèmes et complications rencontrés avec la négation, qui « donnent à la logique pertinente un caractère d'étrangeté sémantique » tout à fait étranger à ses motivations philosophiques :

> « Sa motivation philosophique originelle était de construire un modèle formel du raisonnement dépourvu du principe général, $A, {\sim}A \vdash B$, ou *Ex Falso Quodlibet,* et où consistance simple et consistance absolue soient distinctes. Cette malheureuse connexion entre des idéaux pertinents et le dialéthéisme [*i.e.* l'idée qu'un énoncé peut être à la fois vrai et faux] peut être ramenée dans chaque cas à un autre aspect que ces programmes sémantiques ont à tort en commun : l'explication classique de la validité.
> « Il y a une explication. Bob Meyer a récemment observé que le développement de la sémantique était une réponse aux critiques des logiciens classiques, critiques selon lesquelles les systèmes pertinents manquaient d'une sémantique, et qu'en conséquence ses auteurs avaient fermement entrepris de "prêcher les Gentils dans leur propre langue". Si bien que son critère de validité, ou de conséquence logique, était tout à fait consciemment le critère classique, selon lequel la conjonction de prémisses vraies et d'une conclusion fausse, dont l'impossibilité entraîne la validité, est tout à fait extensionnelle » (Read, 1988).

Tant qu'on conserve en effet la définition classique de la conséquence logique : A entraîne B si et seulement s'il est impossible que A soit vrai et B faux, la modalité étant représentée par une quantification sur les points d'une structure de modèle, l'une des solutions pour éviter qu'une contradiction n'entraîne n'importe quoi est d'admettre qu'elle peut être vraie en certains points. Il y a bien sûr d'autres moyens d'obtenir ce résultat, comme d'admettre parmi les valeurs de vérité la valeur à-la-fois-Vrai-et-Faux. Mais fondamentalement l'objectif est le même. Or cette définition restée classique de la conséquence logique, dans le métalangage, est profondément étrangère à la visée initiale

1. La négation de la logique minimale de Johansson n'obéit pas à *Ex Falso Quodlibet Sequitur.*

des programmes de logique pertinente, qui était de réformer la notion de conséquence *logique*, et de formaliser une nouvelle notion (peut-être *la vraie* notion). Le point de départ en était en effet l'idée que si la préservation de la vérité des prémisses à la conclusion est une condition nécessaire de la validité d'une inférence logique, ce n'en est pas une condition suffisante : quelque chose de plus est requis, qu'on cherche à caractériser. Donc, même si la relation « A entraîne B », telle que formellement définie dans la sémantique de Routley-Meyer, coïncidait indiscutablement en extension avec la relation préthéorique « A entraîne de manière pertinente », via leur commune coextensivité avec « $\mathbf{R} \vdash A \rightarrow B$ », il n'en resterait pas moins que les deux relations diffèrent en intension. Et les problèmes avec la négation sont le résultat de cette distorsion.

La solution : abandonner la définition classique de la conséquence, ou plutôt la rectifier (pour Read, il s'agit évidemment de rectifier une erreur : Read est « pertinentiste »). La conjonction, dans « il est impossible que les prémisses soient vraies *et* la conclusion fausse » doit être comprise, non comme la conjonction extensionnelle classique, mais comme la conjonction intensionnelle, « fusion », ou **o.** La définition authentique de la conséquence logique (pertinente, mais cela va sans dire) est alors la suivante :

A *entraîne* B <u>ssi</u> il est impossible que A soit vrai **o** (fusion) B faux.

Comment cette définition règle-t-elle le problème de EFQ ? On ne veut pas que $(A \wedge {\sim}A)$ entraîne B, il faut donc que $(A \wedge {\sim}A)$ **o** ~ B soit possiblement vrai, sans que cela implique qu'une contradiction soit vraie. Mais, pour cela, il faut évidemment que A **o** B n'entraîne pas A (sinon $(A \wedge {\sim}A)$ **o** ~ B entraînerait la contradiction $(A \wedge {\sim}A)$, ce qu'on veut justement éviter). Mais c'est bien le cas, et c'est l'un des traits par où la conjonction intensionnelle et la conjonction extensionnelle se distinguent le plus nettement : $A \wedge B$ entraîne A (et aussi B), mais A **o** B n'entraîne ni A, ni B. En effet, A **o** B est équivalent à $\sim (A \rightarrow \sim B)$, et peut être défini dans **R** par cette formule. Si A **o** **B** entraînait A, ~A entraînerait $(A \rightarrow {\sim}B)$, ce qui n'est évidemment pas le cas (je reviendrai sur cette remarque un peu plus bas).

Jusqu'ici, le nouveau connecteur **o** a été introduit dans le métalangage, pour obtenir une redéfinition de la notion de conséquence valide. Mais un second argument montre qu'il faut aussi en disposer dans le langage-objet, pour une raison entr'aperçue par Anderson et Belnap (vol. I, § 7 .2). Si l'on veut avoir quelque chose comme la règle →Introduction, disons dans une présentation en séquents, il ne faut pas qu'une collection de prémisses ait toujours le sens d'une conjonction[1]. En effet, si $A, B \vdash C$ permet de dériver le séquent :

$A \vdash B \rightarrow C$,

et si la virgule - disons-le selon une formulation approximative - a le sens d'une conjonction, comme on a :

$A, B \vdash A$,

on peut dériver immédiatement le séquent :

1. Voir le chapitre 7 pour une présentation plus précise des formulations en calcul des séquents.

$A \vdash B \to A,$

dont on ne veut surtout pas. Il faut donc distinguer un autre sens de la ponctuation dans les prémisses, qui correspond au connecteur fusion, et limiter la règle →Introduction à ce type de contexte. Car alors on a bien :

$$\frac{A \mathbf{o} B \vdash C}{A \vdash B \to C,}$$

mais on n'a pas :

$A \mathbf{o} B \vdash A,$

ni en ce nouveau sens de la ponctuation dans la collection de prémisses :

$A, B \vdash A.$

Rétrospectivement, le raisonnement qui menait de $A, B \vdash A$ à $A \vdash B \to A$, peut être compris comme fondé sur une confusion entre les deux connecteurs *et* et *fusion,* ou entre deux sens possibles de la virgule. Pour résumer :

> « Un argument est valide si et seulement s'il est impossible pour les prémisses d'être vraies *fusion* la conclusion fausse. Il ne découle pas de cela qu'une contradiction entraîne n'importe quelle proposition. La formulation correcte de l'explication standard [de la validité] utilise la notion intensionnelle de conjonction, fusion. L'erreur classique consiste à supposer que la combinaison entre prémisses et conclusion qui est à exclure est la simple conjonction extensionnelle.
> « On a aussi besoin de fusion pour la formulation correcte de l'Équivalence de la Déduction [*i.e.* le contenu du théorème de la déduction]. (...) La nécessité de fusion dans la formulation de l'Équivalence de la Déduction montre qu'on a besoin de fusion, non seulement au niveau métalinguistique, pour caractériser correctement la validité, mais aussi au niveau du langage-objet, pour ce qui est de la combinaison des prémisses d'un argument » (Read, 1988).

On pourrait protester contre une certaine circularité, au moins apparente, de l'explication : pour définir la conséquence logique, on fait appel à la conjonction intensionnelle, mais d'autre part, quand il s'agit d'expliquer celle-ci, on fait appel au sens supposé connu de →, censé exprimer une notion de conséquence. A-t-on réellement avancé ? Mais il est préférable de penser que ces explications de **o** par « implique » ne sont que des commentaires préliminaires, à des fins pédagogiques : en fait, la véritable spécification de **o** est à attendre des règles structurelles d'un système formel de déduction (l'étude de ces règles structurelles est renvoyée au prochain chapitre)[1].

Il n'y a donc aucune raison de penser que la négation en logique pertinente, sous prétexte qu'une sémantique conceptuellement mal fondée l'a représentée par une opération qui tolère des contradictions vraies, n'est pas la négation classique habituelle. Pour couper court à l'accusation de style quinien qu'en « changeant de logique, on a changé de sujet », le mieux qu'on puisse faire est de construire une sémantique

1. « Ainsi la différence [entre les logiciens classiques et pertinents] est centrée sur la caractérisation de fusion, distinct de "∧". Une manière de le caractériser est de le lier à l'implication. Dans **R**, "A **o** B" est équivalent à"~(A → ∼B)". Mais ce n'est pas vrai dans des systèmes sans la règle Permutation ; et c'est inutilement circulaire. La meilleure solution est de caractériser fusion, comme nous l'avons fait, relativement à une opération algébrique sur les prémisses, la combinaison-de-prémisses, satisfaisant certaines conditions structurelles plausibles» (Read, 1988).

« homophonique », où les connecteurs du langage-objet sont interprétés par les connecteurs habituels de notre langage familier, utilisé comme métalangage :

> « Nous pouvons éviter cette accusation en construisant une sémantique pour la logique pertinente qui respecte deux traits. Premièrement, elle doit être élaborée dans un métalangage pertinent, dans lequel l'explication de la correction [*soundness*], ou de la conséquence logique, est pertinent - il est impossible pour les prémisses d'être vraies fusion la conclusion fausse. En second lieu, elle doit respecter l'explication homophonique de la signification des connecteurs (...), par exemple :
>
> "si A alors B" est vrai <u>ssi</u> si A est vrai, alors B est vrai,
>
> et : "non-A" est vrai <u>ssi</u> A n'est pas vrai.
>
> [en anglais, de manière plus strictement homophonique : "not-A" is true iff A is not true] » (Read, 1988).

On aura ainsi administré la preuve que la négation qui figure dans la logique pertinente, la négation de De Morgan *en ce sens-là,* reste bien la négation classique utilisée dans le métalangage, notre négation « ordinaire », sans pour autant qu'une contradiction implique n'importe quoi, grâce à la rectification de la notion de conséquence logique.

Le but d'une telle sémantique est de reconstruire dans un métalangage cette fois formalisé, et qui contienne les moyens d'expression adéquats, le concept de conséquence logique pertinente tel qu'il a été caractérisé plus haut, puis de permettre de démontrer un théorème d'adéquation, selon lequel A → B est un théorème de **R** si et seulement si B est conséquence logique de A au sens défini. Je résume l'exposé de cette sémantique de manière relativement informelle.

Le métalangage **MR** pour **R** doit donc contenir les mêmes connecteurs que **R** lui-même, mais cette fois compris comme des symboles conventionnels pour les significations des particules ordinaires : « non », « et », « si ... alors », etc., ou à tout le moins supposés compris, comme dans le cas de **o**. Il contient en outre un prédicat de vérité, « <u>Vr</u>(*x*) », des constantes d'énoncé **p, q, r…,** pour nommer les variables *p, q, r*..., de **R,** et des lettres d'énoncé P, Q, R..., destinées à traduire dans **MR** les variables de **R,** tout cela afin de pouvoir construire une définition de la vérité pour les formules de **R**. Les constantes d'énoncés et les paramètres P, Q, R..., de **MR** sont supposés énumérés, et le schéma d'axiome sémantique ci-dessous spécifie une « traduction » de **R** dans **MR,** de sorte que la lettre d'énoncé X est la traduction de la variable de **R** dont le nom est la constante $\mathbf{x}_i$. On suppose par ailleurs donné un procédé de formation de noms des formules complexes de **R** à partir des constantes d'énoncés (si **a**, **b** sont des noms de formules, on dira que **a** ∧ **b** est un nom de la conjonction de ces formules, etc.).

Axiomes sémantiques :

$\underline{\mathrm{Vr}}(\mathbf{x}_i) \leftrightarrow X_i$

Clauses récursives pour la définition du prédicat de vérité :

$\underline{\mathrm{Vr}}(\sim\mathbf{a}) \leftrightarrow \sim\underline{\mathrm{Vr}}(\mathbf{a})$
$\underline{\mathrm{Vr}}(\mathbf{a} \wedge \mathbf{b}) \leftrightarrow \underline{\mathrm{Vr}}(\mathbf{a}) \wedge \underline{\mathrm{Vr}}(\mathbf{b})$
$\underline{\mathrm{Vr}}(\mathbf{a} \vee \mathbf{b}) \leftrightarrow \underline{\mathrm{Vr}}(\mathbf{a}) \vee \underline{\mathrm{Vr}}(\mathbf{b})$
$\underline{\mathrm{Vr}}(\mathbf{a} \rightarrow \mathbf{b}) \leftrightarrow \underline{\mathrm{Vr}}(\mathbf{a}) \rightarrow \underline{\mathrm{Vr}}(\mathbf{b})$
$\underline{\mathrm{Vr}}(\mathbf{a}\ \mathbf{o}\ \mathbf{b}) \leftrightarrow \underline{\mathrm{Vr}}(\mathbf{a})\ \mathbf{o}\ \underline{\mathrm{Vr}}(\mathbf{b})$.

L'appareil déductif de **MR** est identique à celui choisi pour **R** (système d'axiomes ou déduction naturelle). On étend **MR** par l'ajout d'un opérateur de modalité, ◊, à lire « il est possible que », avec les règles ou les axiomes appropriés (par exemple de type **S4).** On introduit dans **MR**◊ le connecteur ⇒ par la définition : $E \Rightarrow F =_{Df} \sim \lozenge (E\ \mathbf{o} \sim F)$, où E, F sont des formules de **MR**◊, conformément à la proposition de définir la conséquence logique via la conjonction intensionnelle, et on définit un prédicat binaire destiné à exprimer la relation de conséquence logique entre deux formules **a** et **b de R,** <u>Cons</u>(**a**, **b),** par <u>Vr</u>(**a**) ⇒ <u>Vr</u>(**b**). On peut alors démontrer le résultat disant qu'il est vrai (*i.e.* dérivable) dans **MR**◊ qu'une formule **b** est Conséquence de **a** si et seulement si $A \to B$ est un théorème de **R,** où A et B sont les formules de **R** dont les noms sont **a** et **b** respectivement[1] :

MR◊ ⊢ <u>Cons</u>(**a**, **b)** <u>ssi</u> R ⊢ A → B,

autrement dit, la dérivabilité d'une implication dans **R** correspond exactement à la relation de conséquence (pertinente) définie dans le métalangage.

1) On montre d'abord que pour toute formule A de **R** (de nom **a**), il existe une formule A' de **MR** (sa traduction) telle que :

MR ⊢ <u>Vr</u>(**a**) ↔ A'

(la démonstration a lieu par récurrence sur la complexité des formules, à partir des axiomes sémantiques pour les formules atomiques).

2) On montre ensuite que **MR**◊ ⊢ E → F <u>ssi</u> **MR**◊ ⊢ E ⇒ F :

Preuve (en déduction naturelle), de droite à gauche :

~ ◊ (Eo-F)

~ ◊ (E **o** ~F)	théorème, et définition de ⇒
E	hypothèse
~F	hypothèse
E **o** ~F	**o** -Introduction
◊ (E **o** ~F)	◊-Introduction
~ ◊ (E **o** ~F)	Réitération
~~F	~Introduction
F	Double Négation Élimination
E → F	→Introduction

1. Le lecteur sourcilleux aura noté l'abus de langage qui consiste à parler d'une formule **a** de R ; en toute rigueur, il faudrait parler d'une formule de nom **a**.

De gauche à droite :

$E \to F$	théorème
$\quad \Diamond(E \mathbf{o} \sim F)$	hypothèse 1
$\qquad E \mathbf{o} \sim F$	hypothèse 2
$\qquad \sim(E \to F)$	déductible à partir de $(E \mathbf{o} \sim F)$
$\qquad E \to F$	Réitération
$\quad \sim(E \to F)$	$\Diamond$-Élimination sur l'hypothèse 2
$\quad E \to F$	Réitération
$\sim \Diamond (E \mathbf{o} \sim F)$	~Introduction[1]

On peut alors obtenir le résultat désiré : $\mathbf{MR}\Diamond \vdash \underline{\mathrm{Cons}}(\mathbf{a}, \mathbf{b})$ ssi $\mathbf{R} \vdash A \to B$.

Preuve :

	$\mathbf{R} \vdash A \to B$	
ssi	$\mathbf{MR} \vdash A' \to B'$	A' et B' étant les traductions de A et B dans **MR,** et les mêmes dérivations étant admissibles
ssi	$\mathbf{MR}\Diamond \vdash A' \to B'$	**MR**◊ étant une extension conservative de **MR**[2]
ssi	$\mathbf{MR}\Diamond \vdash \underline{\mathrm{Vr}}(\mathbf{a}) \to \underline{\mathrm{Vr}}(\mathbf{b})$	par 1) et théorème de remplacement
ssi	$\mathbf{MR}\Diamond \vdash \underline{\mathrm{Vr}}(\mathbf{a}) \Rightarrow \underline{\mathrm{Vr}}(\mathbf{b})$	par 2)
ssi	$\mathbf{MR}\Diamond \vdash \underline{\mathrm{Cons}}(\mathbf{a}, \mathbf{b})$	par définition de Cons

Quelle est la portée de cette preuve ? En un sens, elle est « clairement circulaire », puisque les ressources déductives de **MR** ne font que reproduire celles de **R** Toute inadéquation de **R** relativement à la « vraie » notion de conséquence se répercuterait donc dans **MR,** qui de ce fait ne peut servir à justifier **R,** ou à convaincre quelqu'un de sa correction. En un autre, elle est parfaitement irréprochable, la définition de la relation de conséquence dans **MR**◊ ne faisant qu'exprimer en termes précis et formels la notion « il est impossible que A soit vrai *fusion* B faux »[3]. Sur la base de cette définition, une sémantique homophonique est possible, qui invalide EFQ et SD sans distordre les conditions de vérité de la négation : la négation de **R,** la négation dite « pertinente » reste donc la négation classique telle qu'elle est dans notre langage d'arrière-fond. De sorte que l'accusation d'avoir changé de sujet en changeant de logique tombe.

L'objection qu'on pourrait faire à l'encontre de toute l'entreprise de Read serait plutôt la suivante : le projet de *définir explicitement* la notion de conséquence pertinente était étranger aux visées originelles d'Anderson et Belnap, qui recherchaient une caractérisation axiomatique de cette notion, fondée sur des motifs plausibles. Read estime que la notion de pertinence est beaucoup trop vague pour permettre d'isoler les inférences valides : partir de là, c'est mettre la charrue avant les bœufs. C'est au

1. La règle utilisée dit que si une formule A est contradictoire avec un théorème, la formule ◊A l'est aussi ; on peut donc décharger A La preuve serait plus simple si le métalangage contenait en outre l'opérateur de nécessité, car la règle de nécessitation (plus l'équivalence de N~ avec ~◊) donnerait immédiatement le résultat recherché.

2. Non démontré ici (la preuve de Read fait appel aux structures de modèles pour **R** de Routley-Meyer enrichies d'un domaine D de variables propositionnelles).

3. . « Que nous apprend donc la Proposition [qu'on vient de démontrer] ? Pour celui dont les préférences logiques sont capturées par **R**, elle explicite de manière formelle la théorie de la vérité et de la conséquence. Ce n'est pas de là qu'on peut partir pour convertir quelqu'un à **R**. Vue comme une défense de la théorie de la conséquence dans **R**, elle est clairement circulaire. Son intérêt consiste à clarifier en termes formels les notions de vérité et de conséquence » (Read, 1988).

contraire la reconnaissance de la validité de certaines inférences qui permet de dire qu'il y a pertinence des prémisses pour la conclusion. Et cette reconnaissance s'appuie sur une définition de la conséquence logique où fusion joue un rôle fondamental (Read, 1988, chap. 6). C'est là que des doutes sont permis : je ne suis pas sûr qu'une définition explicite de la notion de conséquence logique (de la vraie et unique notion : Read, on l'a dit, est « pertinentiste ») à partir du connecteur fusion apporte toutes les lumières qu'en attend Read, tant ce connecteur garde quelque chose d'artificiel : s'agit-il encore d'explication vraiment homophonique de la signification quand on caractérise le connecteur **o** par l'idée de conjonction intensionnelle ? L'introduction de ce connecteur est visiblement une construction théorique, liée en particulier (quoique pas seulement : voir le problème des déductions sous hypothèses) au souci de trouver un motif de l'invalidité d'EFQ formulé avec la négation classique. Mais pourquoi peut-on désirer rejeter EFQ, sinon en raison de la faute de pertinence qu'il exhibe clairement ? Ce n'est certainement pas en raison d'une connaissance préalable de fusion, qui nous montrerait que ce mode d'inférence n'est pas admissible. En bref, la conjonction intensionnelle ne me paraît pas la notion éclairante, primitive, et fondamentale, à partir de laquelle définir, ou redéfinir, la notion de conséquence logique : l'idée de pertinence, aussi indécise soit-elle, reste bien une règle heuristique pour nos évaluations. Si tel est le cas, l'entreprise de Read, quel que soit son intérêt, - montrer que la logique pertinente peut parfaitement s'accommoder d'une interprétation classique de la négation -, n'est pas entièrement convaincante. Mais au point où nous en sommes, ce sont toutes les tentatives de formuler une sémantique pour **R** qui semblent se heurter à des obstacles fondamentaux.

CHANGER DE LOGIQUE, CHANGER DE SUJET ?

De quoi parle-t-on, finalement, quand on parle en logique pertinente de la négation de De Morgan ? Est-ce encore la négation classique, simplement généralisée ? Est-ce une nouvelle négation, avec des propriétés différentes, - mais une négation tout de même, parce que les propriétés *essentielles* de la négation sont préservées ? Ou peut-être n'est-ce même plus une négation ? Le problème qui se pose ici est celui de l'identité de la notion exprimée par un connecteur d'une logique à l'autre, ou, dans les termes de Quine, le problème du « changement de sujet ». À propos de l'idée d'admettre qu'une contradiction soit vraie, tout en évitant l'inconsistance absolue (la déductibilité de tout énoncé) grâce à un changement de logique, Quine écrit ces lignes célèbres :

> « Mon point de vue sur ce dialogue [entre deux logiciens divergeant sur les conséquences d'une contradiction], c'est qu'aucune des deux parties ne sait de quoi elle parle. Elles pensent qu'elles parlent de la négation, "~", "non" ; mais il est sûr que la notation n'est plus identifiable comme la négation quand on se met à considérer des conjonctions de la forme "p. ~p" comme vraies, tout en cessant de considérer que de tels énoncés impliquent tous les autres. Ici, bien évidemment, réside la -difficulté du logicien déviant : quand il tente. de refuser la doctrine, il change seulement de sujet » (Quine, 1970, chap. 6).

Routley et Meyer ont parfois présenté la négation de De Morgan comme une négation classique, simplement « généralisée » à de nouveaux contextes, et on a déjà discuté la vraisemblance de cette affirmation. Mais il y a des arguments directs qui semblent montrer que la négation de De Morgan *n'est pas* la négation classique. L'un d'eux s'appuie sur l'existence du système **KR.** Ce système part d'une remarque de Urquhart (1972), qui avance l'idée qu'il y a en fait deux classes de « paradoxes » différents, que rejette en même temps la logique pertinente. Les paradoxes de la première classe sont bien liés à des fautes de pertinence : ajout d'hypothèses sans lien avec les conclusions dans $A \rightarrow (B \rightarrow A)$, comme dans $A \rightarrow (B \rightarrow B)$. Mais ceux de la seconde reposent sur l'exclusion de points inconsistants, et sont donc plutôt des « paradoxes de consistance » : admettre de tels points permet de rejeter $(A \wedge {\sim}A) \rightarrow B$. D'où l'idée d'explorer des systèmes qui resteraient pertinents en un sens, parce que refusant les formules de la première classe, tout en admettant les formules caractéristiques de la seconde. La logique **KR** est le système **R** auquel a donc été rajouté comme axiome :

$(A \wedge {\sim}A) \rightarrow B$ (EFQ).

Urquhart commente :

> « Ici nous avons un système de logique pertinente avec la négation Booléenne régulière classique. (...) Dans **KR** la négation classique et la négation pertinente sont *identifiées* » (Anderson et Belnap, 1992, § 65.1.2).

Naturellement, il faut modifier la sémantique pour **R** si l'on veut rendre valide cette formule. Une **KR**-structure de modèle est une **R-**structure à laquelle on impose la nouvelle condition :

$a^* = a$, pour tout point a ;

Inversion (si R*abc,* alors R*ac*b**) devient donc : si R*abc,* alors R*acb,* de sorte qu'on a la symétrie totale : si R*abc,* alors R*bac* et R*acb.* Et par l'identification de a avec a^*, la clause pour la négation devient :

$a \vDash {\sim}A$ <u>ssi</u> $a \nvDash A$.

Il semble clair qu'il s'agit à présent de la négation classique, si l'on admet que ce qui caractérise la négation classique, ce n'est pas forcément d'opérer avec une sémantique vérifonctionnelle à deux valeurs de vérité. C'est que A et ~A ne soient jamais vraies ensembles (mais qu'au moins l'une des deux le soit), de quelque façon que la sémantique représente ce « jamais ». Ce fait peut être assuré aussi bien par l'opération de complémentation booléenne sur des treillis, que par une sémantique avec des points indexant les valeurs de vérité, pourvu que l'opération * soit définie par $a = a^*$ pour tout point. Avec cette négation classique, EFQ devient évidemment valide pour cette sémantique. C'est un argument puissant en faveur de l'idée qu'une sémantique, construite dans le même cadre (seules quelques propriétés de R et de * ont changé), mais qui invalide EFQ, ne donne pas à la négation son sens classique.

Bien que l'ajout de EFQ comme axiome ne fasse pas de **KR** une nouvelle formulation de la logique classique, le changement de négation a des effets y compris sur le fragment positif (*i.e.*, sans négation) de la logique pertinente. Par exemple, la formule (parmi d'autres, voir§ 54.1 de Anderson et Belnap, 1992 ; mais elle a l'avantage de la simplicité) :

$(A \mathbf{\ o\ } A) \vee (A \rightarrow B)$

devient valide, alors qu'elle n'est pas **R**-valide. Supposons en effet que (A → B) ne soit pas vérifié en 0 (vrai en 0 pour une valuation). Alors il existe *a* tel que A est vrai en *a*. Par le Postulat R0*aa* et la symétrie totale, on a : R*a*0*a*, et R*aa*0. Donc il existe bien un point *x* tel que R*xx*0 où A est vrai, et donc (A **o** A) est vérifié en 0. La formule ressemble tant à la curieuse tautologie classique A ∨ (A → B) qu'on peut y voir une faute manifeste de pertinence (ce qui, soit dit en passant, laisse planer quelques doutes sur l'étanchéité des deux classes de paradoxes distinguées par Urquhart).

Le second argument est discuté longuement au § 80 d'Anderson et Belnap, 1992, à propos d'une extension possible de **R** suggérée par Meyer et Routley. Au lieu d'ajouter EFQ comme axiome, on introduit à côté de ~ une nouvelle négation, dite booléenne, notée ¬, pour laquelle la clause d'évaluation est la clause familière :

a ⊨ ¬A ssi *non a* ⊨ A.

(C'est à dessein que *non* est souligné). Et naturellement, cette nouvelle négation a la propriété que (A ∧ ¬A) → B et (¬A ∧ (A ∨ B)) → B sont valides. Puisque les formules analogues mais où figure la négation de De Morgan ne sont pas valides, il semble qu'on ait là un nouvel argument en faveur de la thèse qu'on a bien « changé de sujet ». Mais les auteurs s'efforcent de montrer que les choses ne sont pas aussi simples. Je reprends l'essentiel de leur mise en scène du débat entre le « pertinentiste » (*i.e.*, celui qui croit fermement que la notion de conséquence tautologique, définie pour le système des conséquences du premier degré, donc pour les connecteurs classiques, est la seule notion correcte de conséquence entre fonctions de vérité), et le « classicaliste », qui ne croit qu'aux canons classiques de l'inférence. Les auteurs reconnaissent que le pertinentiste, au point de départ (à la fois historiquement et conceptuellement) doit tenir que la négation qui figure dans le syllogisme disjonctif (ou dans EFQ) est bien la négation classique :

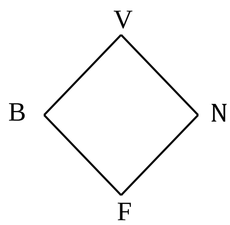
> « Le pertinentiste orthodoxe, ou le logicien pertinent jouant le rôle du pertinentiste, a une explication de l'interaction entre négation et conséquence, qui a l'allure suivante. En logique, nous trouvons les connecteurs vérifonctionnels ordinaires, en particulier la conjonction, la disjonction, et la négation, et ce que la logique classique nous dit à leur sujet est en partie vrai, et en partie faux. Ce qui est vrai, c'est ce qu'elle nous dit au sujet de la question : quels énoncés ne comportant que les fonctions de vérité sont logiquement vrais. Ce qui est faux, c'est ce qu'elle nous dit concernant la question : pour ce genre d'énoncés, lesquels entraînent quels autres. (...) Le pertinentiste présume qu'il parle des mêmes connecteurs vérifonctionnels que le logicien classique, mais insiste sur le fait qu'une relation différente, plus étroite, doit être prise comme relation de conséquence » (Anderson et Belnap, 1992, § 80.2.2).

Rappelons, en effet, que c'est là la position qu'il semblait qu'il faille adopter pour répondre valablement au « défi de Geach » : c'est bien du syllogisme disjonctif classique, avec sa négation classique, dont il est question, et dont on veut nier la validité. Mais à présent qu'on a introduit, à côté de la négation de De Morgan, une négation visiblement classique, il semble qu'on ne puisse plus soutenir la thèse initiale, selon laquelle la première négation était aussi classique.

> « Or, maintenant que l'existence d'une négation booléenne est notée par le classicaliste, il peut bien répondre que les pertinentistes font plus qu'insister simplement sur le fait qu'il doit prendre "conséquence" en un sens non classique - ils insistent aussi sur le fait qu'il doit prendre "négation" en un sens non classique. La négation classique, c'est la négation

> Booléenne, et EFQ et le SD valent pour la négation booléenne (...). Et le classicaliste peut continuer en faisant remarquer que, même si un pertinentiste parvenait à montrer, par un argument habile, que la négation de De Morgan est après tout la négation "réelle", peut-être bien après tout la même que la "négation classique", la "négation" booléenne serait toujours là » (*ibid.*).

Ce qui est la substance des arguments proposés plus haut, en faveur de la thèse que, vue du point de vue de la sémantique relationnelle, la négation de De Morgan *n'est pas* la négation classique. Mais les auteurs de *Entailment* prétendent que le pertinentiste peut refuser à bon droit d'entendre l'argument - pas seulement de l'accepter, mais même de le comprendre. Comment cela ? Plaçons nous, pour les besoins de la discussion, dans le cadre d'une sémantique à quatre valeurs de vérité, Vrai (V), Faux (F), Vrai-et-Faux (B), et ni-Vrai-ni-Faux (N), qui sont ordonnées de manière à former le treillis suivant (souvenons-nous que la discussion porte sur la notion de conséquence tautologique, avec sa sémantique des treillis de De Morgan) :



Interprétons ces quatre valeurs de vérité comme les quatre parties de l'ensemble $\{v, f\}$, *i.e.* $V = \{v\}$, $F = \{f\}$, $B = \{v, f\}$, et $N = \emptyset$, et notons [A] la valeur d'une formule quelconque A. La négation ~ est interprétée par une opération (notée de la même façon, le contexte dissipant toute ambiguïté) telle que :

$$\sim V = F, \sim F = V, \sim B = B, \sim N = N.$$

On reconnaît dans ce treillis « logique » un exemple de treillis de De Morgan (à quatre éléments), ce qui justifie l'idée que ~ ainsi interprété est la négation de De Morgan. Cette interprétation est intuitivement justifiée : si une formule a la valeur B, sa négation a aussi la valeur B (fausse puisque la première est vraie, mais aussi vraie puisque la première est fausse). Et si une formule a la valeur N, sa négation aussi, dans la mesure où l'on ne sait rien de la première. À présent, l'autre négation $\neg$ est interprétée par une nouvelle opération (notée là encore identiquement $\neg$), semblable à la négation de De Morgan sur V et F, mais telle que $\neg B = N$, et $\neg N = B$. Il est clair que, pour cette opération, le treillis est complémenté, ou booléen. Là encore, une justification plausible est la suivante : si une formule est à la fois vraie et fausse, sa négation ne peut être vraie (il faudrait pour cela que la première formule ne soit pas vraie), et ne peut être fausse non plus pour une raison analogue. Les règles d'évaluation pour les deux négations peuvent alors être présentées de la manière suivante :

De Morgan	$(v\sim)$	$v \in [\sim A]$ <u>ssi</u> $f \in [A]$
	$(f\sim)$	$f \in [\sim A]$ <u>ssi</u> $v \in [A]$
Booléenne	$(v\neg)$	$v \in [\neg A]$ <u>ssi</u> non $(v \in [A])$
	$(f\neg)$	$f \in [\neg A]$ <u>ssi</u> non $(f \in [A])$.

Question : comment doit-on comprendre le « non » du métalangage qui figure dans les clauses pour la négation booléenne ? La réponse paraît s'imposer : de manière pertinente, c'est-à-dire (du moins pour les auteurs) comme une négation de De Morgan : « Mais le vrai pertinentiste doit utiliser pour lui-même un métalangage pertinent, avec une seule négation, qui soit la négation de De Morgan » (*ibid.*). Mais, en ce point, on se trouve confronté à la difficulté suivante : si le métalangage est interprété dans la sémantique à quatre valeurs de vérité, et si le « non » qui figure dans les clauses est compris comme une négation de De Morgan, alors il n'y a plus moyen de distinguer les deux négations, et l'argument de la distinction tombe.

Remplaçons, dans les deux clauses pour la négation booléenne, « non » par ~, cette fois en tant que symbole du métalangage. Admettons par ailleurs deux principes sémantiques minimaux concernant la valeur des énoncés métalinguistiques disant que v (ou f) est au nombre des valeurs de A :

$$[v \in [A]] = [A] \qquad \text{et} \qquad [[f \in [A]] = [\sim A].$$

On a alors :

$[f \in [A]] = [\sim A]$	principe sémantique
$[\sim A] = \sim [A]$	interprétation de ~
$\sim[A] = \sim[v \in [A]]$	principe sémantique
$\sim[v \in [A]] = [\sim(v \in [A])]$	interprétation de ~

et par transitivité :

$$[f \in [A])] = [\sim (v \in [A])]$$

d'où (°) :

$f \in [A])$ <u>ssi</u> $\sim (v \in [A])$.

En clair, le Faux est au nombre des valeurs de A si et seulement si le Vrai n'en est pas, ce qui revient, semble-t-il, à exclure la possibilité qu'à une même formule soit assigné à la fois Vrai-et-Faux, ou ni l'un ni l'autre : à exclure donc une sémantique à quatre valeurs de vérité. De plus :

	$v \in [\sim A]$ <u>ssi</u> $f \in [A]$	clause (v~)
	j $\in [A]$ <u>ssi</u> $\sim (v \in [A])$	par (°)
d'où :	$v \in [\sim A]$ <u>ssi</u> $\sim (v \in [A])$	

ce qui est exactement la clause ($v\neg$) dès lors que « non » est interprété comme la négation de De Morgan :

$v \in [\neg A]$ <u>ssi</u> $\sim (v \in [A])$

d'où finalement :

$v \in [\sim A]$ <u>ssi</u> $v \in [\neg A]$.

On obtient de même :

$f \in [\sim A]$ <u>ssi</u> $f \in [\neg A]$,

de sorte que moyennant le principe sémantique, on a : $[\sim A] = [\neg A]$, et les deux négations se confondent. Ce résultat montre d'abord qu'un métalangage pertinent, *i.e.*

utilisant la négation de De Morgan, ne permet pas de formuler les conditions de vérité de la négation booléenne de manière à la distinguer de la négation de De Morgan. C'est un résultat curieux. L'explication suggérée est qu'il est impossible de formuler une sémantique à quatre valeurs de vérité dans un métalangage pertinent. Il est par exemple impossible, dans un métalangage contenant la négation de De Morgan, de dire qu'un énoncé prend juste la valeur *v* (ou juste la valeur *f*). On pourrait croire que c'est pourtant ce qu'exprime un énoncé comme :

$$v \in [\mathrm{A}] \wedge \sim (f \in [\mathrm{A}]),$$

qui semble dire que le Vrai est au nombre des valeurs de A, mais pas le Faux. Mais c'est oublier qu'avec la négation de De Morgan du métalangage, le second membre de la conjonction ne dit pas que le Faux *n'est pas* parmi les valeurs de A. D'après le principe sémantique incorporé dans la théorie, la valeur de $(f \in [\mathrm{A}])$ est identique à la valeur de ~A, de sorte que $\sim (f \in [\mathrm{A}])$ est équivalent à ~~A, c'est-à-dire à A, qui a même valeur que $(v \in [\mathrm{A}])$. Autrement dit, le second membre de la conjonction ne dit rien de plus que le premier : quelque chose comme « le Faux n'est pas l'unique valeur de A », ce qui laisse ouverte la possibilité que A ait les deux valeurs à la fois.

Si l'on admet les deux principes sémantiques minimaux, la preuve de la confusion des deux négations semble incontestable. Elle est cependant affectée d'une certaine circularité : pour couper court à l'argument des deux négations différentes, l'argumentation a supposé que le métalangage était lui-même, et devait être, plutôt que classique, pertinent. Jusque-là, tout va bien : un vrai « pertinentiste » n'a pas à s'accommoder d'un métalangage classique. Mais du coup, la preuve présuppose qu'un métalangage pertinent est nécessairement un langage où la négation est la négation de De Morgan. Autrement dit, la distinction dont elle montre qu'elle est inassignable dans le langage objet est présupposée dans le métalangage, comme caractéristique d'un métalangage pertinent. C'est précisément ce point qui est discutable. Le mérite de l'analyse de Read était de montrer qu'un langage pertinent ne requiert nullement une nouvelle négation, non classique.

La moralité finale de toutes ces réflexions semble être que si la sémantique de Routley-Meyer représente la négation par la négation de De Morgan, cette représentation n'est nullement indispensable à l'analyse de la négation en logique pertinente. Les chapitres suivants ont pour but, en particulier, d'explorer ce que le point de vue *proof-theoretic* peut nous apprendre sur la négation pertinente.

7. LE POINT DE VUE STRUCTURAL

> « Les pertinentistes ont suivi la tradition classique consistant à se focaliser sur l'ensemble des énoncés valides plutôt que sur ce qui est l'aspect réellement important d'une logique : la (les) relation(s) de conséquence qui lui sont associées. »
>
> Avron, 1992.

QU'EST-CE QU'UNE COLLECTION DE PRÉMISSES ?

Je voudrais pour commencer éclairer le diagnostic d'« échec partiel » que posait Avron dans son article « Whither Relevant Logic? » (Avron, 1992), sur l'approche d'Anderson et Belnap. Brutalement formulé, le problème est le suivant : A est-il déductible des deux prémisses A, B ? Les auteurs ne fournissent pas de réponse claire à cette question.

L'idée de base, on l'a vu, était d'introduire dans le langage-objet un connecteur d'implication destiné à représenter la relation de déductibilité : A → B si et seulement si A *déduit* B (si l'on permet cette expression pour « B est déductible de A »). Du coup, se pose immédiatement la question de savoir dans quelles circonstances A déduit réellement B, c'est-à-dire dans quels cas il y a réellement une preuve de B à partir de A. Et, en fait, plusieurs concepts de preuve sont proposés dans *Entailment*, on l'a vu en particulier à propos de la règle Adjonction. Or le problème majeur naît du fait qu'usuellement la relation de déductibilité n'est pas une relation entre deux formules, mais une relation entre une collection de formules, les prémisses, et (disons pour l'instant) une formule, la conclusion. La notion de conséquence « à la Tarski » (*Tarski-like*) consigne ce fait. Et donc, si l'on veut que le connecteur d'implication représente la relation de déductibilité, il semble, *dans le cadre qui était celui d'Anderson et Belnap*, qu'il faille d'abord réduire le cas général (collection de formules ⊢ formule) au cas particulier d'une relation entre deux formules (comme remarqué par Avron dans l'article cité). Comment opérer cette réduction ? On peut être guidé sur cette voie par deux intuitions, qui n'aboutissent pas du tout aux mêmes résultats.

1) La première idée qu'on peut avoir est de réduire une collection de prémisses à la conjonction de ces prémisses : $A_1, ..., A_n$ déduit B <u>ssi</u> ($A_1 \wedge ... \wedge A_n$) déduit B. C'est ce qu'Avron appelle le *Principe de Réduction* : une collection de prémisses a la valeur d'une conjonction. Avec ce principe, il n'est pas nécessaire que toutes les prémisses A_i aient été utilisées dans la déduction de B, puisque, pour que la conjonction des prémisses déduise B, il suffit que *certaines* des prémisses soient pertinentes pour B, il n'est pas exigé que toutes le soient. Comme l'écrivaient Anderson et Belnap à propos d'un des axiomes pour la conjonction :

« La conséquence A&B → A n'est pas infectée par la présence de B, qui est possiblement sans pertinence : comme déjà remarqué, c'est là précisément le genre de non-pertinence que la notion purement extensionnelle, vérifonctionnelle, de conjonction est destinée à traiter » (Anderson et Belnap, 1975, § 23.6).

Par voie de conséquence, sous le Principe de Réduction, il n'est pas nécessaire que tous les A_i soient pertinents pour B pour que $A_1, ..., A_n$ déduise B.

2) La seconde idée qui peut se présenter est toute différente. Elle met l'accent sur le lien étroit de déductibilité entre (disons) A_n et B *dans le contexte des autres prémisses* $A_1, ..., A_{n-1}$ c'est-à-dire sur l'exigence que A_n ait été effectivement utilisé dans la déduction de B pour conclure à la conséquence $A_n \rightarrow B$ à partir des autres prémisses. Par exemple :

« En d'autres termes, ce qu'on recherche est un système (...) pour lequel il y a un théorème de la déduction prouvable à l'effet qu'il existe une preuve de B *à partir* [souligné : *i.e.* où B est utilisé] de l'hypothèse A si et seulement si A → B est prouvable » (*ibid.*, § 3).

Cette intuition est aussi fidèle à l'idée que A entraîne B ssi B est déductible de A, mais elle accommode autrement le contexte des autres prémisses. C'est cette idée qui mène au théorème de la déduction pertinent, et par itération, à l'idée que :

$$A_1, ..., A_n \text{ déduit B } \underline{\text{ssi}} \vdash A_1 \rightarrow (... \rightarrow (A_n \rightarrow B) ...).$$

On a là bien sûr une autre notion de déduction, où toutes les prémisses doivent être pertinentes pour B[1]. Cela montre clairement que la notion de déductibilité utilisée de manière heuristique par Anderson et Belnap a du mal à affronter les situations ordinaires où il s'agit de collections de prémisses. En bref : une telle collection est-elle à comprendre comme une conjonction, ou comme une suite de conséquences emboîtées ?

En fait, malgré les obscurités d'*Entailment* sur ce point, le Principe de Réduction (au sens où Avron utilise ce terme) n'était pas entièrement admis par Anderson et Belnap, et ne pouvait pas l'être. On n'en a que la moitié, dans le sens : s'il y a une preuve (dans **E** ou dans **R**, et pour un certain concept de preuve qui n'exige pas que toutes les prémisses soient utilisées) que $A_1, ..., A_n$ déduit B, alors $(A_1 \wedge ... \wedge A_n) \rightarrow B$ est un théorème (de **E** ou de **R**), et donc B est déductible de $(A_1 \wedge ... \wedge A_n)$. C'est le contenu du « théorème de la conséquence » (§ 23 .6). Mais on n'a pas l'implication dans l'autre sens : si $(A_1 \wedge ... \wedge A_n)$ déduit B, alors $A_1, ..., A_n$ déduit B. Car $A \wedge B$ déduit évidemment $A \wedge B$, et si on en concluait qu'il y a une preuve de $A \wedge B$ à partir des deux prémisses A, B, on aurait une preuve de $A \rightarrow (B \rightarrow (A \wedge B))$, apparemment pertinente, puisque A et B sont utilisées :

1. Les deux relations de déductibilité sont parfois opposées comme déductibilité externe *versus* déductibilité interne (voir par exemple Avron, 1999).

$$\dfrac{\dfrac{\dfrac{A, B}{A \wedge B}}{B \to (A \wedge B)}}{A \to (B \to (A \wedge B))}$$

En fait, accepter ce sens du Principe de Réduction reviendrait à admettre la formule en question (d'où, on l'a vu, la restriction sur les indices dans les applications de ∧Introduction).

Anderson et Belnap ont reconnu que l'option 1), consistant à interpréter une collection de prémisses comme une conjonction, ne pouvait être retenue, sauf à abandonner ou limiter de quelque façon l'idée sous-jacente au théorème de la déduction (comme on le voit dans le théorème de la conséquence, où les prémisses sont en un sens déchargées « en bloc »). En témoignent par exemple ces lignes :

> « Il est clair (...) que la suite de formules à gauche du signe de déduction (*turnstile*) ne peut avoir le sens d'une conjonction de prémisses. Car, selon l'interprétation conjonctive, nous aurions $A, B \vdash A$, et la règle en question [*i.e.* →Introduction] mènerait de cela à $A \vdash B \to A$, et delà à $\vdash A \to (B \to A)$, qui n'est pas bon » (Anderson et Belnap, § 7. 2).

Mais si une collection n'a pas (toujours ?) le sens d'une conjonction, comment doit-on la comprendre ? Comme des emboîtements d'implications, répondent de manière plutôt énigmatique Anderson et Belnap (§ 7. 2). Les prémisses impliquent la conclusion « in a *nested* sense » :

$A_1, ..., A_n$ déduit B <u>ssi</u> $\vdash A_1 \to (... \to (A_n \to B))$,

bien qu'ils reconnaissent que ce n'est pas ainsi que nous comprenons ordinairement une collection de prémisses (§ 22. 2. 2 par exemple). L'option 2) manque de clarté. De plus, cette incertitude sur ce qui peut valoir comme preuve conduit à multiplier les concepts de déduction. Comme le note Avron : « Il est d'une certaine manière surprenant qu'il n'y ait pas exactement une définition généralement acceptée, officielle, de ce qu'est une preuve à partir d'assomptions dans la masse des travaux pertinentistes » (Avron, 1992).

Ces remarques préliminaires avaient pour but de justifier l'étude directe de la relation de déductibilité et de ses propriétés. On va voir que le point de vue structural permet de clarifier grandement les deux manières de comprendre ce qu'est une collection de prémisses.

Il est usuel de considérer la relation de déductibilité comme une relation entre *ensembles* d'énoncés (où de formules, dans un langage formel) et énoncés (ou formules), relation notée :

$\Gamma \vdash C$.

Mais le point de vue de Gentzen dans la présentation du Calcul des séquents était, en un sens, d'emblée plus général. Les antécédents des séquents y sont conçus comme des *suites* finies de formules, de sorte qu'une notation plus appropriée est :

$A_1, ..., A_n \vdash C$,

qu'on peut lire : de la suite $A_1, ..., A_n$, la formule C est déductible. De plus, dans le Calcul classique, dont c'est la caractéristique, on admet des suites de formules également dans les conséquents. Mais alors, comme le note Gentzen, la virgule n'a pas le même sens dans l'antécédent et le conséquent : elle a la valeur d'une conjonction dans l'antécédent, mais la valeur d'une disjonction dans le conséquent. Cet aspect contextuel du signe de ponctuation suggérait à lui seul qu'une plus grande attention devait être portée aux différentes manières de concevoir ce qu'est une collection (pour employer le terme le plus neutre) de formules.

Limitons-nous pour l'instant aux collections de formules dans les antécédents. On peut se demander en quoi la considération de suites de formules est plus générale que la considération d'ensembles, alors que les premiers objets sont plus complexes que les seconds. Mais c'est précisément cette plus grande complexité qui permet à Gentzen de rendre explicites certaines règles, dites « règles structurales », qui ont pour effet de réduire la structure des suites à la structure d'objets moins complexes. Par exemple, si l'on part d'une suite de formules, où donc l'ordre des éléments compte, et qu'on applique une règle de Permutation qui permet d'intervertir les éléments de la suite, cela revient à considérer que l'ordre est indifférent pour la relation de déductibilité, ou, en d'autres termes, à identifier une collection de prémisses avec un ensemble de prémisses, comportant éventuellement des répétitions (puisque, dans la suite initiale, le même élément pouvait figurer à différentes places). À partir d'un tel « multi-ensemble », si une règle de Contraction permet de supprimer des répétitions de formules à volonté, on peut obtenir finalement un objet qui a les propriétés d'un ensemble. En bref, plus les collections de prémisses sont conçues comme initialement structurées, plus nombreuses sont les règles de déduction qu'on peut expliciter, et qui sont susceptibles de réduire cette complexité. Et naturellement, une fois que ces règles sont explicitées, il est possible de les refuser, soit dans un but purement exploratoire, soit pour des raisons conceptuellement motivées. Le rejet, ou la restriction à certaines situations, de telle ou telle de ces règles donne lieu à la pluralité des logiques dites substructurales :

> « Nous proposons d'appeler les logiques qu'on peut obtenir de cette manière, en restreignant les règles structurales, *logiques substructurales.* La logique dont nous restreignons les règles structurales afin d'obtenir les logiques substructurales est en principe la logique classique » (Dosen, 1993).

Cette proposition de baptême a eu gain de cause, et l'appellation est devenue standard. Cependant, si on est sensible au geste théorique qui a permis d'expliciter les règles qu'on peut éventuellement rejeter, à savoir considérer primitivement des collections *structurées* de prémisses, pour *ensuite* formuler des règles réduisant plus ou moins cette complexité, ces systèmes de logique mériteraient aussi bien le nom de *logiques structurales.* Les présentations données ci-dessous de différents systèmes font justement grand usage de cette notion de *structures de prémisses.* En ce sens, c'est une généralisation des idées de Gentzen. Il faut convenir, cependant, que la règle Affaiblissement, qui permet d'ajouter de nouvelles prémisses, n'entre pas exactement dans ce cadre : elle est plutôt solidaire d'une propriété plausible de la relation de conséquence, la monotonie, *i.e.* :

Si $\Gamma \vdash C$, alors $\Gamma ; A \vdash C$;

mais l'admission d'Affaiblissement a des effets sur les propriétés des structures de prémisses.

À des fins de clarté, on commence par présenter différents systèmes en déduction naturelle sous forme de séquents[1].

SYSTÈMES POSITIFS EN DÉDUCTION NATURELLE

On considère d'abord des langages propositionnels contenant les connecteurs binaires $\rightarrow$, **o**, $\wedge$, $\vee$, et un symbole compté parfois comme un connecteur à zéro places ***t***, la constante du (logiquement) vrai. La notion de formule est définie comme d'ordinaire (à l'absence près de la négation), ***t*** étant compté comme une formule. La nouveauté consiste dans l'introduction de signes de ponctuation, ayant, comme les connecteurs, une arité définie (à la différence de la virgule de Gentzen) : deux signes de ponctuation binaires : **;** et **,** , et un signe de ponctuation à zéro place, 0. Ces signes permettent de définir inductivement la notion de *structure*[2] :

1) une formule est une structure ;

2) si P est un signe de ponctuation, d'arité zéro ou deux, le résultat de l'application de P à zéro ou deux structures (selon le nombre d'arguments), est une structure (par exemple, 0, X**;**Y, Z**,** (X**;**Y), etc., sont des structures, la convention étant d'écrire les signes de ponctuation binaires entre les deux arguments et de parenthéser le tout).

De même que les formules contiennent des sous-formules, les structures contiennent des sous-structures. La notion est définie ainsi :

1) toute structure est sous-structure d'elle-même ;

2) si une structure Z est de la forme PXY, les sous-structures de X et de Y sont sous-structures de Z.

La notion de substitution est fréquemment utilisée dans la suite : on note X(Y) une structure contenant Y comme sous-structure, et X(Z) la structure obtenue en remplaçant cette occurrence de Y par la structure Z dans X.

Un séquent est une figure de la forme X $\vdash$ A, où X est une structure, l'antécédent, et A une formule, le conséquent. Une inférence est une paire, dont le premier membre est un ensemble de zéro, un, ou deux séquents prémisses, et le second membre est un séquent, la conclusion. S'il y a zéro prémisse, l'inférence est un *axiome*. Une preuve d'un séquent est constituée par un arbre, dont la racine est le séquent à prouver, les sommets des axiomes, et où chaque pas (nœud) correspond à une application d'une règle d'inférence (on peut élargir cette définition à des preuves avec assomptions, en acceptant que figurent aux sommets, outre des axiomes, ces assomptions).

La justification du remplacement de la notion de suite de formules par celle de structuration binaire n'est pas immédiate, mais apparaîtra clairement dès lors que les règles structurales, qui ont pour effet de réduire (en un certain sens) des collections à d'autres collections, seront explicitées. C'est dans l'article « Display Logic » que Belnap

1. Pour une présentation en séquents de la Déduction naturelle classique, voir par exemple Sundholm, 1983.

2. Je m'inspire largement pour toute cette partie de l'exposition très claire de Greg Restall (Restall, 2000).

a introduit les signes de ponctuation sous le nom de « connecteurs de structures », en justifiant ainsi l'innovation[1] :

> « Le seul moyen qui ne laisse passer aucune possibilité est de prendre une structuration binaire, plutôt que polyvalente. Et je pense, à la réflexion, que cela est plus dans l'esprit de Gentzen (...).
> « Je ne prétends pas que la structuration binaire est plus intuitive, mais bien qu'elle est plus satisfaisante d'un point de vue mathématique. Je recommande la structuration binaire pour les mêmes raisons qui conduisent presque tout le monde à préférer dans les systèmes formels la conjonction binaire à une conjonction polyvalente » (Belnap, 1982).

Le schéma d'axiome suivant, ainsi que les règles pour les connecteurs, sont communs à tous les systèmes construits sur la base du langage décrit plus haut. Les règles structurales, en revanche, sont « en libre service », au sens où l'on peut ou non les adopter ; la présence (ou l'absence) de telle ou telle règle caractérise un système de logique déterminé.

1) Un seul schéma d'axiome (règle à zéro prémisse), où A est n'importe quelle formule :

$$A \vdash A.$$

2) Règles opératoires (une règle d'introduction et une règle d'élimination, pour chaque connecteur ; X, Y, etc., sont des structures) :

$$\frac{X;A \vdash B}{X \vdash A \to B}\ \to\mathrm{I} \qquad\qquad \frac{X \vdash A \to B \qquad Y \vdash A}{X;Y \vdash B}\ \to\mathrm{E}$$

$$\frac{X \vdash A \qquad Y \vdash B}{X;Y \vdash A\ \mathbf{o}\ B}\ \mathbf{o}\mathrm{I} \qquad\qquad \frac{X \vdash A\ \mathbf{o}\ B \qquad Y(A;B) \vdash C}{Y(X) \vdash C}\ \mathbf{o}\mathrm{E}$$

$$\frac{X \vdash A \qquad X \vdash B}{X \vdash A \wedge B}\ \wedge\mathrm{I} \qquad\qquad \frac{X \vdash A \wedge B}{X \vdash A} \qquad \frac{X \vdash A \wedge B}{X \vdash B}\ \wedge\mathrm{E}$$

$$\frac{X \vdash A}{X \vdash A \vee B} \qquad \frac{X \vdash B}{X \vdash A \vee B}\ \vee\mathrm{I} \qquad\qquad \frac{X \vdash A \vee B \qquad Y(A) \vdash C \qquad Y(B) \vdash C}{Y(X) \vdash C}\ \vee\mathrm{E}$$

$$0 \vdash \boldsymbol{t}\quad \boldsymbol{t}\mathrm{I} \qquad\qquad \frac{X \vdash \boldsymbol{t} \qquad Y(0) \vdash C}{Y(X) \vdash C}\ \boldsymbol{t}\mathrm{E}$$

1. En fait, l'idée de distinguer des suites extensionnelles de prémisses, et des suites intensionnelles, est déjà présente dans la présentation du Calcul des Séquents pour **R+** par Dunn (Anderson et Belnap, 1975, § 28.5).

3) Règles structurales

– Règles pour le signe de ponctuation 0 (à gauche et à droite) :

$$\frac{X \vdash C}{0;X \vdash C}\ {+0} \qquad \frac{0;X \vdash C}{X \vdash C}\ {-0}$$

$$\frac{X \vdash C}{X;0 \vdash C}\ {0+} \qquad \frac{X;0 \vdash C}{X \vdash C}\ {0-}$$

Dans les présentations à la Gentzen, avec des suites de formules dans l'antécédent, on admet des séquents de la forme ⊢ A, où l'antécédent est la suite vide de formules : un tel séquent prouvable correspond à un théorème. Mais du point de vue structural, il n'y a pas de « structure vide », et 0 tient le rôle de la suite vide chez Gentzen. Les séquents qui correspondent à des théorèmes sont donc de la forme « 0 ⊢ A ».

Règles structurales possibles pour **;** (il est naturel d'admettre que non seulement la structure de l'antécédent pris en bloc peut être remaniée, mais que des sous-structures peuvent l'être à l'intérieur de structures. Dans les antécédents de la forme V(X) ci-dessous, V représente un contexte quelconque ; qui peut être une structure ou le signe vide, auquel cas l'antécédent se réduit à X) :

Associativité **(B)** :

$$\frac{V\,(X;(Y;Z)) \vdash C}{V\,((X;Y);Z) \vdash C}$$

Associativité converse **(Be) :** la règle **(B)** lue de bas en haut.

Associativité avec permutation **(B')** :

$$\frac{V\,(X;(Y;Z)) \vdash C}{V\,((Y\,;X)\,;Z) \vdash C}$$

Commutativité forte : **(C)**

$$\frac{V\,((X;Y);Z) \vdash C}{V\,((X;Z);Y) \vdash C}$$

Commutativité faible :

$$\frac{V(X;Y) \vdash C}{V\,(Y;X) \vdash C}$$

Contraction forte : (**W**)

$$\frac{V\,((X;Y);Y) \vdash C}{V(X;Y) \vdash C}$$

Contraction faible :

$$\frac{V(X;X) \vdash C}{V(X) \vdash C}$$

Affaiblissement : (**K**)

$$\frac{V(X) \vdash C}{V(X;Y) \vdash C}$$

Affaiblissement à gauche[1] :

$$\frac{V(X) \vdash C}{V\,(Y;X) \vdash C}$$

Exemples :

1) Preuve de $0 \vdash A \to (B \to (A\ \mathbf{o}\ B))$:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A \qquad B \vdash B}{A;B \vdash A\ \mathbf{o}\ B}\ \mathbf{o}\mathrm{I}}{A \vdash B \to (A\ \mathbf{o}\ B)}\ \to\mathrm{I}}{0;\ A \vdash B \to (A\ \mathbf{o}\ B)}\ +0}{0 \vdash A \to (B \to (A\ \mathbf{o}\ B))}\ \to\mathrm{I}$$

Ce séquent est donc prouvable dans tout système où la règle +0 (introduction de 0 à gauche) est admise, indépendamment des autres règles structurales.

2) Preuve de $(A\ \mathbf{o}\ B) \to C \vdash A \to (B \to C)$:

1. Les noms des règles (**B**) (**C**) (**W**) et (**K**) sont les noms traditionnels ; ils proviennent de la logique combinatoire de Cuny, où (**B**) est le « combinateur » qui représente la composition de deux fonctions, (**C**) la transformation d'une fonction en la fonction converse, (**W**) la fonction à un argument obtenue à partir d'une fonction à deux arguments par identification des deux arguments, et (**K**) l'expression d'une constante comme fonction d'une variable (Guny et Feys, 1958).

$$\dfrac{\dfrac{\dfrac{\dfrac{(A\ \mathbf{o}\ B) \to C \vdash (A\ \mathbf{o}\ B) \to C \qquad \dfrac{A \vdash A \qquad B \vdash B}{A;B \vdash A\ \mathbf{o}\ B}\ \mathbf{o}\mathrm{I}}{(A\ \mathbf{o}\ B) \to C;(A;B) \vdash C}\ {\to}\mathrm{E}}{((A\ \mathbf{o}\ B) \to C;A);B \vdash C}\ \text{Associativité}}{(A\ \mathbf{o}\ B) \to C;A \vdash (B \to C)}\ {\to}\mathrm{I}}{(A\ \mathbf{o}\ B) \to C \vdash A \to (B \to C)}\ {\to}\mathrm{I}$$

3) Preuve de A → (B → C) ⊢ (A **o** B) → C:

$$\dfrac{\dfrac{A\ \mathbf{o}\ B \vdash A\ \mathbf{o}\ B \qquad \dfrac{B \vdash B \qquad \dfrac{A \to (B \to C) \vdash A \to (B \to C) \qquad A \vdash A}{A \to (B \to C);A \vdash (B \to C)}\ {\to}\mathrm{E}}{(A \to (B \to C);\ A);B \vdash C}\ {\to}\mathrm{E}}{(A \to (B \to C);A\ \mathbf{o}\ B \vdash C}\ \text{Associativité et } \mathbf{o}\mathrm{E}}{A \to (B \to C) \vdash (A\ \mathbf{o}\ B) \to C}\ {\to}\mathrm{I}$$

4) Preuve de ***t*** → A ⊢ A :

$$\dfrac{\dfrac{\boldsymbol{t} \to A \vdash \boldsymbol{t} \to A \qquad 0 \vdash \boldsymbol{t}}{\boldsymbol{t} \to A;0 \vdash A}\ {\to}\mathrm{E}}{\boldsymbol{t} \to A \vdash A}\ 0–$$

Ce séquent est à comparer avec l'axiome du système **E**+ : ((A → A) → B) → B. Si l'on interprète ***t*** comme le logiquement vrai, il dit que si une formule est conséquence d'une vérité logique (ou de la conjonction de toutes les vérités logiques), elle est vraie, ce qui paraît assez raisonnable. Ainsi présenté, **E+** admet donc la règle structurale 0– (élimination de 0 à droite). En revanche, il n'admet pas la règle 0+, qui permet de dériver le séquent A ⊢ ***t*** → A.

5) Preuve de A ⊢ ***t*** → A :

$$\dfrac{\dfrac{\boldsymbol{t} \vdash \boldsymbol{t} \qquad \dfrac{A \vdash A}{A;0 \vdash A}\ 0+}{A;\boldsymbol{t} \vdash A}\ \boldsymbol{t}\mathrm{E}}{A \vdash \boldsymbol{t} \to A}\ {\to}\mathrm{I}$$

Sous la même interprétation de ***t***, ce séquent correspond en effet à l'un des théorèmes typiques de **R+**, Démodalisation : il dit en effet que si une formule A est vraie, alors elle est conséquence du logiquement vrai, donc nécessaire, ce que rejette **E**.

6) Preuve de (A → B) ∧ (A → C) ⊢ (A → (B ∧ C)). Cette preuve ne met en jeu aucune règle structurale, et le séquent est donc admissible dans tous les systèmes quel que soit le choix des règles (la preuve n'est qu'à moitié écrite, l'autre partie étant identique, au remplacement près des occurrences soulignées de B par C) :

$$\dfrac{\dfrac{\dfrac{(A\to B)\wedge(A\to C)\vdash(A\to B)\wedge(A\to C)}{(A\to B)\wedge(A\to C)\vdash(A\to \underline{B})}\;\wedge E \qquad A\vdash A}{\big((A\to B)\wedge(A\to C)\big)\mathbf{;}\ A\vdash \underline{B}}\;\to E}{\big((A\to B)\wedge(A\to C)\big)\mathbf{;}\ A\vdash B\wedge C}\;\wedge I$$

7) Si la règle d'associativité (**Bc**) est admise pour **;** , le connecteur **o** est lui aussi associatif, *i.e.* :

$$A\ \mathbf{o}\ (B\ \mathbf{o}\ C)\vdash(A\ \mathbf{o}\ B)\ \mathbf{o}\ C$$

Preuve :

$$\dfrac{A\ \mathbf{o}\ (B\ \mathbf{o}\ C)\vdash A\ \mathbf{o}\ (B\ \mathbf{o}\ C) \qquad \dfrac{B\ \mathbf{o}\ C\vdash B\ \mathbf{o}\ C \qquad \dfrac{\dfrac{\dfrac{A\vdash A \qquad B\vdash B}{A\mathbf{;}B\vdash A\ \mathbf{o}\ B}\;\mathbf{o}I \qquad C\vdash C}{(A\mathbf{;}B)\mathbf{;}C\vdash(A\ \mathbf{o}\ B)\ \mathbf{o}\ C}\;\mathbf{o}I}{A\mathbf{;}(B\mathbf{;}C)\vdash(A\ \mathbf{o}\ B)\ \mathbf{o}\ C}\;(\mathbf{Bc})}{A\mathbf{;}(B\ \mathbf{o}\ C)+(A\ \mathbf{o}\ B)\ \mathbf{o}\ C}\;\mathbf{o}E}{A\ \mathbf{o}\ (B\ \mathbf{o}\ C)\vdash(A\ \mathbf{o}\ B)\ \mathbf{o}\ C}\;\mathbf{o}E$$

8) Avec Commutativité faible, on obtient le séquent correspondant à la Loi d'assertion, *i.e.* :

$$0\vdash A\to\big((A\to B)\to B\big)$$

Preuve :

$$\dfrac{\dfrac{\dfrac{A\to B\vdash A\to B \qquad A\vdash A}{A\to B\mathbf{;}A\vdash B}\;\to E}{A\mathbf{;}A\to B\vdash .B}\;(\mathbf{C})\text{ faible}}{0\mathbf{;}A\vdash(A\to B)\to B}\;\to I$$

Une dernière étape permet d'obtenir le séquent recherché.

9) Avec les règles (**K**) et Affaiblissement à gauche, on obtient finalement les séquents caractéristiques de la logique classique :

$$0\vdash A\to(B\to A) \quad \text{et} \quad 0\vdash A\to(B\to B).$$

Preuves :

$$\dfrac{\dfrac{A\vdash A}{A\mathbf{;}\ B\vdash A}}{A\vdash B\to A} \qquad\qquad \dfrac{\dfrac{B\vdash B}{A\mathbf{;}\ B\vdash B}}{A\vdash B\to B}$$

Dans ce cadre, on peut spécifier un système de logique par le vocabulaire retenu, c'est-à-dire le choix des connecteurs, et les règles structurales admises, et l'on peut convenir de nommer un système en prenant la liste des règles qui le déterminent. Le fragment positif de **E**, **E**+ (sans la distributivité sur laquelle on reviendra) correspond ainsi au système dont le langage contient tous les connecteurs introduits jusqu'ici ; et dont les règles sont : 0-introduction et élimination à gauche, 0-élimination à droite, Associativité et (**B'**), Contraction. Autrement dit, le système +0–00–**BB'W**, l'absence de mention des connecteurs indiquant qu'ils sont tous présents. L'absence de 0-introduction à droite se justifie par le fait, vu plus haut, que cette règle permet de prouver le séquent $A \vdash \boldsymbol{t} \to A$, qui contrevient à la prise en compte par **E** de la nécessité ; l'absence de la règle (**C**), par le fait qu'elle permet de prouver le séquent qui correspond à Permutation, que **E** n'admet pas sans restriction :

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A \to (B \to C) \vdash A \to (B \to C) \qquad A \vdash A}{A \to (B \to C); A \vdash B \to C} \qquad B \vdash B}{(A \to (B \to C); A); B \vdash C}}{(A \to (B \to C); B); A \vdash C}\ \text{Règle }(\mathbf{C})}{A \to (B \to C); B \vdash A \to C}}{A \to (B \to C) \vdash B \to (A \to C)}$$

Le système +0–00+0–**BCW** (avec tous les connecteurs) correspond au fragment positif (sans la distributivité) de **R**, **R**+. En fait, en présence de Commutativité, les deux règles de 0-introduction et élimination à gauche (ou à droite) suffisent. Le système **BC** (sans Contraction ni Affaiblissement) donne un fragment de la logique linéaire positive, qui utilise par ailleurs d'autres connecteurs. Le système **BCWK,** qui possède toutes les règles structurales, et dont le langage est enrichi d'un connecteur .l, muni de la règle du faux :

$$\frac{X \vdash \bot}{Y(X) \vdash A} \qquad \bot E$$

correspond à la logique intuitionniste, la négation de A étant définie par $A \to \bot$. En l'absence de la constante du faux, on a la logique minimale.

LA DISTRIBUTIVITÉ

On peut admettre que la conjonction extensionnelle ne soit pas distribuée par la disjonction (extensionnelle). On peut se souvenir ici que, sémantiquement parlant, tous les treillis considérés à titre d'interprétations de $\mathbf{R}_{\mathbf{fde}}$ (chap. 4) n'étaient pas nécessairement distributifs ; et que, du point de vue syntaxique, la distributivité avait dû être ajoutée comme une règle primitive au système de déduction naturelle du chapitre 5 (voir Anderson et Belnap, 1975, § 23.3). Certains des systèmes présentés ici admettent la possibilité que la distributivité ne soit pas démontrable, puisqu'il n'est pas possible de prouver le séquent $A \wedge (B \vee C) \vdash (A \wedge B) \vee (A \wedge C)$ sans utiliser les deux formes d'Affaiblissement, et Contraction faible, qu'on peut rejeter. Par exemple, **R+** ainsi caractérisé n'admet pas la distributivité.

Cependant, on peut aussi décider de « vivre avec la distributivité »[1]. C'est à cette fin qu'a été introduit le signe de ponctuation **,** , non utilisé jusqu'ici, pour lequel les règles Associativité, Commutativité, Contraction et Affaiblissement valent (plus exactement, les versions de ces règles avec **,** remplaçant **;**). En voici une preuve faisant usage du signe de ponctuation extensionnel, ainsi qualifié parce qu'il a les mêmes propriétés que la conjonction extensionnelle (de même que **o** mime le comportement de **;**). La preuve est disposée de manière linéaire par manque de place :

1	$A \wedge (B \vee C) \vdash A \wedge (B \vee C)$	
2	$A \wedge (B \vee C) \vdash A$	∧E sur 1
3	$A \wedge (B \vee C) \vdash B \vee C$	∧E sur 1
4	$A \wedge (B \vee C), B \vdash A$	Affaiblissement pour **,** sur 2
5	$B \vdash B$	
6	$A \wedge (B \vee C), B \vdash B$	Affaiblissement pour **,** sur 5
7	$A \wedge (B \vee C), B \vdash A \wedge B$	∧I sur 4 et 6
8	$A \wedge (B \vee C), B \vdash (A \wedge B) \vee (A \wedge C)$	∨I sur 8
9	$C \vdash C$	
10	$A \wedge (B \vee C), C \vdash C$	Affaiblissement pour **,** sur 9
11	$A \wedge (B \vee C), C \vdash A$	Affaiblissement pour **,** sur 2
12	$A \wedge (B \vee C), C \vdash A \wedge C$	∧I sur 10 et 11
13	$A \wedge (B \vee C), C \vdash (A \wedge B) \vee (A \wedge C)$	∨I sur 12
14	$A \wedge (B \vee C), A \wedge (B \vee C) \vdash (A \wedge B) \vee (A \wedge C)$	∨E sur 3, 8 et 13
15	$A \wedge (B \vee C) \vdash (A \wedge B) \vee (A \wedge C)$	Contraction sur 14

LA RÉGLE DE COUPURE EST ADMISSIBLE

Gentzen ajoute aux règles structurales une règle de Coupure (dite aussi Transitivité), qui dit que dans certaines conditions une formule intermédiaire A, figurant à la fois dans le conséquent et dans l'antécédent de deux séquents prémisses, peut être éliminée, et disparaît au cours d'une déduction :

$$\frac{X \vdash A \qquad Y(A) \vdash B}{Y(X) \vdash B}$$

Puis Gentzen prouve un théorème dit d' « élimination », montrant que toute preuve faisant usage de la règle peut être transformée en preuve « sans coupure ». Dans l'esprit d'Anderson et Belnap, 1975 (§ 7.2), on peut présenter les choses autrement : les systèmes construits ici ne possèdent pas initialement la règle de Coupure, néanmoins celle-ci est *admissible,* bien que non prouvable à partir des règles primitives, au sens où elle n'augmente pas le stock des théorèmes. Ce qui peut être prouvé avec son aide est aussi prouvable sans elle. Voici une preuve rapide de son admissibilité.

1. Pour parodier le titre de l'article de Belnap, « Life in the Undistributed Middle » *(in* Schroeder-Heister et Dosen, 1993).

Définition 1 : Étant donné une inférence tombant sous une règle, des occurrences d'une structure dans cette inférence forment une *famille de paramètres,* si et seulement si, pour toute structure Y, l'inférence obtenue en remplaçant ces occurrences par Y est encore une instance de la même règle.

Définition 2 : Une structure figurant dans une famille de paramètres dans une inférence est un paramètre.

Par exemple, dans une instance particulière de :

$$\frac{X \vdash A \rightarrow B \qquad Y \vdash A}{X;Y \vdash B}$$

X, Y, et leurs sous-structures sont des paramètres, mais ni A, ni B, ni $A \rightarrow B$.

Définition 3 : Deux structures dans une inférence sont congruentes si et seulement si elles appartiennent à la même famille de paramètres. La relation de congruence est évidemment une relation d'équivalence.

Remarque : dans tous les systèmes construits avec l'une quelconque des règles ci-dessus (sauf les axiomes), pour toute inférence, toutes les formules qui figurent dans l'antécédent de la conclusion de cette inférence sont des paramètres (propriété de *régularité de l'antécédent*). Il suffit pour s'en convaincre de regarder les règles : aucune nouvelle formule déterminée n'apparaît dans l'antécédent d'une conclusion qui ne figurât pas déjà dans l'antécédent de la (ou des) prémisse(s). La seule exception apparente est la règle d'Affaiblissement (si on l'adopte), mais elle introduit des formules sans s'occuper de leurs particularités.

PROPOSITION. — La règle de Coupure est admissible dans tous les systèmes de Déduction naturelle construits avec les règles ci-dessus.

Soit une preuve de $Y(A) \vdash B$, étant donné par ailleurs une preuve de $X \vdash A$. On remplace toutes les occurrences congruentes à A dans la preuve de $Y(A) \vdash B$ par X. La nouvelle preuve obtenue est *presque* une preuve de $Y(X) \vdash B$, dans la mesure où chaque inférence (sauf éventuellement les séquents initiaux de la forme $A \vdash A$) reste une application de la même règle que dans la preuve initiale, les règles étant closes pour la substitution dans une famille de paramètres. Quant aux séquents initiaux de la forme $A \vdash A$, ils ont à présent la forme $X \vdash A$. Mais, par hypothèse, nous avons une preuve de $X \vdash A$, qu'il suffit de « coller » à la construction précédente, pour avoir une preuve du séquent $Y(X) \vdash B$. La règle de Coupure est donc admissible dans tous les systèmes présentés dans ce cadre général.

CALCULS DES SÉQUENTS À UNE SEULE CONCLUSION

Ici encore, les séquents sont de la forme $X \vdash A$, où X est une structure et A une formule. Les règles structurales (au choix, comme plus haut), ainsi que les règles d'introduction des connecteurs restent les mêmes. La différence est que la symétrie : règles d'introduction, règles d'élimination, est remplacée par une autre. On a des règles d'introduction à droite de $\vdash$, comme en Déduction naturelle, mais aussi des règles d'introduction à gauche. Comme les règles d'introduction à droite sont les mêmes que dans la version DN, on ne les donnera pas à nouveau. Les règles d'introduction à gauche sont présentées dans le même ordre que les règles d'élimination correspondantes en DN :

$$\frac{X \vdash A \qquad Y(B) \vdash C}{Y(A \rightarrow B\mathbf{;}X) \vdash C} \quad \rightarrow\vdash$$

$$\frac{X(A\mathbf{;}B). \vdash C}{X(A \;\mathbf{o}\; B) \vdash C} \quad \mathbf{o}\vdash$$

$$\frac{X\,(A) \vdash C}{X(A \wedge B) \vdash C} \qquad \frac{X\,(B) \vdash C}{X(A \wedge B) \vdash C} \quad \wedge\vdash$$

$$\frac{X\,(A) \vdash C \qquad X\,(B) \vdash C}{X\,(A \vee B) \vdash C} \quad \vee\vdash$$

$$\frac{X\,(0) \vdash C}{X(\boldsymbol{t}) \vdash C} \quad \boldsymbol{t}\vdash$$

Enfin la règle de Coupure est également admise (sans justification ici), afin de faciliter la comparaison entre les deux types de présentation.

Exemples de dérivations :

Dans tous les systèmes construits en Calcul des séquents sur ces langages, les séquents suivants sont dérivables :

· $A \wedge B \vdash A$ (la preuve est immédiate)

$(A \rightarrow B)\; \mathbf{o}\; A \vdash B$

Preuve:

$$\cfrac{\cfrac{A \vdash A \qquad B \vdash B}{(A \rightarrow B)\mathbf{;}\ A \vdash B}\ \rightarrow\vdash}{(A \rightarrow B)\; \mathbf{o}\; A \vdash B}\ \mathbf{o}\vdash$$

Dans tous les systèmes où la règle Associativité (**B**) est admise, le séquent suivant est dérivable :

. $(A \;\mathbf{o}\; B) \rightarrow C \vdash A \rightarrow (B \rightarrow C)$

Preuve:

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{A \vdash A \qquad B \vdash B}{A;B \vdash A\ \mathbf{o}\ B} \qquad C \vdash C}{(A\ \mathbf{o}\ B) \to C;(A;B)) \vdash C}\ \to\vdash}{((A\ \mathbf{o}\ B) \to C;A);B \vdash C}\ (\mathbf{B})}{(A\ \mathbf{o}\ B) \to C;A \vdash B \to C}}{(A\ \mathbf{o}\ B) \to C \vdash A \to (B \to C)}$$

Disons qu'étant donné un système **S** en DN, le système correspondant **LS** en Calcul des séquents est le système déterminé par la présence des mêmes connecteurs, et des mêmes règles structurales. On peut montrer facilement qu'un séquent X ⊢ A est dérivable dans un système **S** si et seulement s'il est· aussi dérivable dans le système en Calcul des séquents correspondant **LS.**

On le montre ici pour le système +0–00+0–**BCW** = **R+** (sans la distributivité).

Dans un sens : si X ⊢ A est dérivable dans **R+**, il l'est aussi dans **LR+** avec Coupure.

1) Si X ⊢ A est un axiome (règle d'inférence sans prémisses), il est soit de la forme A ⊢ A, soit 0 ⊢ ***t***, et il est dérivable en Calcul des séquents.

2) Si X ⊢ A provient de l'application d'une règle d'introduction ou d'une règle structurelle de **R+**, et si ses prémisses sont dérivables, X ⊢ A l'est aussi dans **LR+** puisque ces règles sont les mêmes dans les deux versions.

3) Restent donc les cas où X ⊢ A provient de l'application d'une règle d'élimination. Il faut montrer que l'effet de ces règles peut être obtenu dans **LR+** en les passant en revue une à une.

3 *a*) X**;**Y ⊢ C provient de →Élimination à partir des deux prémisses X ⊢ A → C, et Y ⊢ A. Supposons que ces deux prémisses aient été dérivées dans **LR+**.

Par **o**-Introduction, on a donc :

$$X;Y \vdash (A \to C)\ \mathbf{o}\ A.$$

Comme (A → C) **o** A ⊢ C (voir le premier exemple donné plus haut), on a par la règle de Coupure dans **LR+** :

$$\frac{X;Y \vdash (A \to C)\ \mathbf{o}\ A \qquad (A \to C)\ \mathbf{o}\ A \vdash C}{X;Y \vdash C}$$

qui est donc dérivable dans **LR+**.

3 *b*) Y(X) ⊢ C provient des deux prémisses X ⊢ A **o** B et Y(A;B) ⊢ C par **o**-Élimination. Supposons ces deux prémisses dérivées dans **LR+**. À partir de la seconde, on peut dériver :

$$Y(A\ \mathbf{o}\ B) \vdash C,$$

et par la règle de Coupure :

$$\frac{X \vdash A \mathbf{o} B \qquad Y(A \mathbf{o} B) \vdash C}{Y(X) \vdash C}$$

3 *c*) X ⊢ A (ou B) provient par ∧Élimination de X ⊢ A ∧ B, qu'on peut supposer dérivé dans **LR+**. Comme A ∧ B ⊢ A est dérivable, on a par la règle de Coupure :

$$\frac{X \vdash A \wedge B \qquad A \wedge B \vdash A}{X \vdash A}$$

3 *d*) Y(X) ⊢ C provient des prémisses X ⊢ A ∨ B, Y(A) ⊢ C, Y(B) ⊢ C par ∨Élimination. Supposons que ces trois prémisses aient été dérivées dans **LR+**. Les deux dernières permettent de dériver par Introduction à gauche :

$$Y (A \vee B) \vdash C,$$

et on a donc par la règle de Coupure :

$$\frac{X \vdash A \vee B \qquad Y (A \vee B) \vdash C}{Y(X) \vdash C}$$

3 *e*) Y(X) ⊢ C provient des prémisses X ⊢ ***t***, et Y(0) ⊢ C, qu'on peut supposer dérivées dans **LR+.** Par ***t***⊢, on a une dérivation de Y(***t***) ⊢ C dans **LR+**, et une application de la règle de Coupure donne Y(X) ⊢ C.

En·bref, des Introductions appropriées à gauche, suivies d'une application de·la règle de Coupure, permettent d'obtenir les mêmes résultats que les règles d'Élimination correspondantes en DN.

On montre à présent dans l'autre sens que si un séquent X ⊢ A est dérivable dans **LR+**, il l'est aussi dans la version en DN de **R+**. Là encore, les axiomes, règles structurales, et règles d'introduction à droite sont les mêmes. Il suffit donc de montrer que le travail des règles d'introduction à gauche est effectué en DN par les règles d'élimination.

1) Si on a une preuve de Y (A → B;X) ⊢ C à partir des prémisses X ⊢ A, et Y(B) ⊢ C, qu'on peut supposer dérivées dans **R+,** on peut construire une preuve en DN par application de la règle de Coupure :

$$\frac{\dfrac{A \to B \vdash A \to B \qquad X \vdash A}{A \to B\mathbf{;}X \vdash B}\ \to\text{Élim} \qquad Y(B) \vdash C}{Y (A \to B\mathbf{;}X) \vdash C}$$

2) Si on a une preuve de X (A **o** B) ⊢ C à partir de X(A**;**B) ⊢ C, on construit en DN la preuve correspondante :

$$\frac{A \mathbf{o} B \vdash A \mathbf{o} B \qquad X(A\mathbf{;}B) \vdash C}{X (A \mathbf{o} B) \vdash C}$$

3), 4), et 5) pour ∧, ∨, et ***t*** sont laissés au lecteur.

L'équivalence déductive entre les formulations en DN et les versions correspondantes en Calcul des séquents se généralise immédiatement aux systèmes logiques autres que **R+**, puisque chaque fois les seules différences concernent la présence ou l'absence des règles d'élimination et d'introduction à gauche.

(Remarque : si on introduit pour **LR+** le signe de ponctuation, pour lequel valent en particulier les règles d'Affaiblissement et de Contraction, on a comme dans **R+** une preuve, quelque peu fastidieuse, de Distributivité).

Le Calcul des séquents intuitionniste (**LI**) peut être construit sur la base d'un langage où figurent les connecteurs $\to$, $\wedge$, $\vee$, avec leurs règles, la constante du faux $\bot$, le signe de ponctuation **,** (pour lequel toutes les règles structurales sont admises), la structure 0, et la constante ***t***. Les règles pour 0 et ***t*** sont les règles précédentes :

$0 \vdash \boldsymbol{t}$ (axiome) $\qquad \dfrac{X(0) \vdash C}{X(\boldsymbol{t}) \vdash C}$

La constante du faux est munie de la règle caractéristique ($\bot$Élimination, ou règle du faux intuitionniste) :

$$\frac{X \vdash \bot}{X \vdash A}$$

Un exemple de dérivation dans cette version de **LI** est par exemple :

$$\dfrac{\dfrac{A \vdash A \qquad \bot \vdash \bot}{A\mathbf{,}\ A \to \bot \vdash \bot}}{A\mathbf{,}\ A \to \bot \vdash B}$$

Si la négation est introduite par $\neg A = A \to \bot$, on a une forme de *Ex Falso Quodlibet.*

Cette version a cependant l'inconvénient de ne pas respecter le « principe de la sous-formule » (selon lequel aucune formule ne disparaît au cours d'une preuve), puisque dans une application de $\bot$Élimination, la constante du faux disparaît. Il est donc plus usuel de prendre la négation comme connecteur primitif, munie des deux règles suivantes, et d'admettre des séquents avec des conséquents vides (plus rigoureusement : des séquents dont les conséquents sont des ensembles à au plus une formule comme élément). On a donc les trois règles suivantes :

$$\frac{X \vdash A}{X\mathbf{,}\ \neg A \vdash} \qquad \frac{X\mathbf{,}\ A \vdash}{X \vdash \neg A} \qquad \frac{X \vdash}{X \vdash B}$$

où le conséquent vide joue le rôle de la constante du faux (les règles pour la négation sont naturelles, si l'on repense à l'équivalence de $\neg A$ avec $A \to \bot$). La dérivation écrite plus haut devient alors :

$$\dfrac{\dfrac{A \vdash A}{A\mathbf{,}\ \neg A \vdash}}{A\mathbf{,}\ \neg A \vdash B}$$

Exemple : Une preuve (sans coupure) d'une forme de Contraposition admissible du point de vue intuitionniste : $A \to \neg B \vdash B \to \neg A$

$$\dfrac{\dfrac{\dfrac{A \vdash A \qquad \dfrac{B \vdash B}{B, \neg B \vdash}}{B, (A \to \neg B, A) \vdash}}{A \to \neg B, B \vdash \neg A}}{A \to \neg B \vdash B \to \neg A}$$

CALCULS DES SÉQUENTS À CONCLUSION MULTIPLE

Les Calculs des séquents peuvent être rendus complètement symétriques si l'on permet que des structures de complexité quelconque puissent figurer dans les conséquents (c'est ainsi que Gentzen distinguait le Calcul classique du Calcul intuitionniste présenté ci-dessus : par l'absence de la restriction à une seule formule dans le conséquent). Le signe de ponctuation 0 peut figurer seul dans un antécédent ou un conséquent, pour jouer le rôle que joue $\emptyset$ quand les collections de prémisses sont conçues comme des ensembles. On ne donne pas d'emblée, à présent, la totalité des règles pour les connecteurs. Pour l'instant, nos langages ne sont pas spécifiés, et peut-être ne disposent même pas de connecteurs. On va voir justement que les connecteurs peuvent être systématiquement introduits à titre de « reflets », à l'intérieur des langages-objets, des signes de ponctuation et des propriétés des structures[1].

On introduit un nouveau signe de ponctuation unaire, °, opérant sur ces structures que sont les formules, permettant de déplacer les formules de part et d'autre de $\vdash$, tout en indiquant leur « statut » indépendamment de la position qu'elles occupent en fait. Dans $X;A \vdash B$, par exemple, A est une partie de l'antécédent. L'opération ° permet de passer de là au séquent $X \vdash A°;B$, où A figure à présent dans le conséquent, tout en indiquant que A est toujours partie antécédente, en un sens à préciser.

Disons qu'une formule est positive si elle est suivie d'un nombre pair de °, négative si elle est suivie d'un ·nombre impair de ° (A est donc positive, A° négative). Disons en outre qu'une formule est partie antécédente d'un séquent <u>ssi</u> elle est composante positive de l'antécédent, ou composante négative du conséquent ; et partie conséquente d'une séquent <u>ssi</u> elle est composante négative de l'antécédent ou composante positive du conséquent. L'opération ° a les propriétés suivantes :

1. Ce point de vue sur les connecteurs a été développé en particulier par Dov Gabbay, dans « A General Theory of Structured Consequence Relations » (Schroeder-Heister et Dosen, 1993).

$$\frac{X;A \vdash Y}{X \vdash A;Y} \qquad \frac{X \vdash A;Y}{X;A^\circ \vdash Y}$$

Elle permet donc de faire figurer une partie antécédente aussi bien dans le conséquent que dans l'antécédent, ou une partie conséquente (dans la seconde règle) aussi bien dans l'antécédent que dans le conséquent. Si on lit ces propriétés de la déduction également dans l'autre sens, on a des équivalences, qui disent qu'une partie antécédente (ou conséquente) peut figurer explicitement dans l'antécédent (ou le conséquent)[1].

Il faut noter que jusqu'ici, l'opérateur ° (le *shift operator*) fonctionne comme un signe de ponctuation opérant sur certaines structures. L'image d'une formule par ° doit se trouver simplement de l'autre côté du signe de déductibilité. On peut à présent ajouter au langage-objet un nouveau connecteur, au sens usuel (qui appliqué à une formule donne une formule) qui récupère les propriétés de l'opération °. On introduit donc le connecteur ¬ avec des règles analogues :

$$\frac{X;A \vdash Y}{X \vdash \neg A;Y}\ (\vdash\neg) \qquad \frac{X \vdash A;Y}{X;\neg A \vdash y}\ (\neg\vdash)$$

Quelles vont être les propriétés de ce connecteur ? Non seulement on peut dériver le séquent $A \vdash \neg\neg A$ (c'était déjà le cas avec la version intuitionniste), mais on peut à présent dériver Double Négation Élimination (ce qui justifie le choix du symbole de négation pour ce connecteur involutif) :

$$\cfrac{\cfrac{A \vdash A}{0 \vdash \neg A;A}\ (\vdash\neg)}{\neg\neg A \vdash A}$$

On peut également dériver une des formes classiquement admissibles de Contraposition :

$$\cfrac{\cfrac{\cfrac{\cfrac{A \vdash A}{0 \vdash A;\neg A} \qquad B \vdash B}{\neg A \to B \vdash B;A}}{\neg A \to B;\neg B \vdash A}}{\neg A \to B \vdash \neg B \to A}$$

Ici encore, Associativité et Permutation ont été tacitement utilisées. Naturellement, la forme de Contraposition dérivée plus haut dans le cadre du Calcul intuitionniste reste dérivable. Avec cette négation involutive, on retrouve donc la négation dite de « De Morgan », caractérisée par les deux axiomes correspondants (voir **R→¬**).

1. Ce sont presque les « Display-equivalences » de Belnap 1982, à ceci près que le signe de ponctuation ° pouvait porter dans cet article sur des structures quelconques. J'ai simplifié la présentation en m'inspirant de la notion de modalité négative de Restall, 2000.

L'idée de comprendre le connecteur de négation comme une traduction dans le langage-objet d'une opération structurale peut aisément être généralisée à d'autres connecteurs. On voit facilement que le connecteur d'implication, avec ses deux règles dont le contenu est équivalent au Modus Ponens et au théorème de la déduction « encode de l'information au sujet de la relation de conséquence dans le langage des propositions » (Restall, 2000) : cette idée est au point de départ de toute la logique pertinente.

$$\frac{X \vdash Y(A) \qquad X'(B) \vdash Y'}{X'(X;A \rightarrow B) \vdash Y(Y')} \rightarrow\vdash \qquad\qquad \frac{X;A \vdash B;Y}{X \vdash A \rightarrow B;Y} \vdash\rightarrow$$

Il est peut-être moins clair en Calcul des séquents qu'il ne l'est en DN avec la règle d'élimination, que la règle de →Introduction à gauche exprime le Modus Ponens. Mais on peut raisonner ainsi pour le voir ; dans la formulation avec une seule formule dans le conséquent, les choses sont plus claires :

$$\frac{X \vdash A \qquad B \vdash C}{(A \rightarrow B;X) \vdash C} \rightarrow\vdash$$

si X permet de déduire A, alors X avec A → B donne B ; et si par ailleurs B implique C, alors par transitivité X avec A → B permet finalement de déduire C. Sous une forme ou une autre, les règles assurent que la relation de déductibilité a lieu entre A et la formule B → C exactement quand B ajouté à A déduit C : la formule B → C « dit » donc que le contexte A étant donné, C est déductible de B. Il n'est donc pas étonnant que les différentes formes d'implication selon les différents systèmes héritent leurs caractéristiques des propriétés variées que peut posséder la relation de déductibilité. À leur tour, ces propriétés dépendant des assomptions faites concernant les signes de ponctuation, c'est-à-dire des règles structurales admises ou refusées.

Il est particulièrement clair dans le cas du connecteur **o** (*fusion*) que son comportement reproduit directement celui du signe de ponctuation **;** sur les structures. La règle **o**-Introduction à gauche assure que A **o** B peut remplacer A**;**B dans l'antécédent :

$$\frac{X\,(A;B) \vdash Y}{X\,(A\ \mathbf{o}\ B) \vdash Y}$$

Il n'y a pas de règle qui assure directement la dérivabilité dans l'autre sens, mais si on relit la règle de **o**-Élimination (en DN) comme une sorte de « règle de coupure » portant sur A **o** B et A;B :

$$\frac{X \vdash A\ \mathbf{o}\ B \qquad Y(A;B) \vdash C}{Y(X) \vdash C}$$

la règle revient à identifier la formule A **o** B et la structure A**;**B. On peut dire aussi : si A **o** B, elle-même impliquée par X, implique A**;**B, qui lui-même implique C, alors par transitivité X implique C. En bref, A **o** B est la formule qui correspond à la structure A**;**B. La conséquence en est que les propriétés du connecteur fusion procèdent des propriétés accordées à l'opération **;** sur les structures (on l'a vu plus haut dans le cas de l'associativité).

Le connecteur de conjonction extensionnelle ∧ a évidemment les propriété du signe de ponctuation **,** si on en dispose. Il est intéressant de voir que les deux connecteurs **o** et ∧ se confondent, si les règles Affaiblissement et Contraction (et éventuellement Permutation) sont admises pour le signe **;** (dans le cadre du Calcul des séquents avec structures à gauche et à droite, elles peuvent être admises des deux côtés). Les deux règles d'Introduction à gauche pour ∧ sont :

$$\frac{X(A) \vdash Y}{X(A \wedge B) \vdash Y} \qquad\qquad \frac{X(B) \vdash Y}{X(A \wedge B) \vdash Y}$$

Si Affaiblissements à gauche et à droite (**K** et **K')** sont admis pour **;**, on a :

$$\frac{X\,(A) \vdash Y}{X\,(A;B) \vdash Y} \qquad\qquad \frac{X\,(B) \vdash Y}{X\,(A;B) \vdash Y}$$

d'où également par **o-**Introduction :

$$X\,(A\ \mathbf{o}\ B) \vdash Y.$$

Les deux règles d'Introduction à droite pour **o** et pour ∧ se distinguent par le fait que la conjonction extensionnelle est « sensible au contexte », *i.e.* que les antécédents des deux conjoints à réunir et les structures où ils figurent doivent être identiques ; avec **o,** ils peuvent différer et sont simplement réunis :

$$\frac{X \vdash Y(A) \qquad X \vdash Y\,(B)}{X \vdash Y(A \wedge B)} \quad \vdash\wedge$$

$$\frac{X \vdash Y(A) \qquad X' \vdash Y'(B)}{X;X' \vdash (Y;A\ \mathbf{o}\ B);Y'} \quad \vdash\mathbf{o}$$

Mais, là encore, si Affaiblissement et Contraction sont autorisés (outre Associativité et Permutation), on peut identifier les structures différentes. On obtient en effet :

$$\frac{X;X' \vdash Y(A);Y' \qquad X;X' \vdash Y;Y'(B)}{X;X' \vdash (Y;Y')\,(A\ \mathbf{o}\ B)}$$

comme pour la conjonction. La différence entre les deux connecteurs provient donc de l'absence des deux règles structurales, et s'efface si elles sont réintroduites.

Puisque l'implication a été introduite à l'aide du signe de ponctuation **;**, et que le connecteur **o** hérite des propriétés de ce dernier, on doit retrouver les relations entre **;** et → sous forme de relations entre **o** et l'implication. Et tel est bien le cas ; on a en effet :

$$\frac{A \mathbf{o} B \vdash C}{A \vdash B \to C} \qquad \text{et} \qquad \frac{A \vdash B \to C}{A \mathbf{o} B \vdash C}$$

comme le lecteur peut s'en assurer.

Les caractéristiques de l'implication dépendent des assomptions concernant les règles structurales. L'exemple type est bien sûr l'effet d'Affaiblissement :

$$\cfrac{\cfrac{\cfrac{A \vdash A}{A\mathbf{;}B \vdash A}}{A \vdash B \to A}}{0 \vdash A \to (B \to A)} \qquad \mathbf{K}$$

i.e. le séquent correspondant à la formule typique des fautes de pertinence.

Considérons à présent la disjonction extensionnelle, dont il n'a pas été question jusqu'ici. Les deux règles d'introduction à droite pour ∨ sont :

$$\frac{X \vdash Y(A)}{X \vdash Y\,(A \vee B)} \qquad \frac{X \vdash Y(B)}{X \vdash Y\,(A \vee B)}$$

Dans un Calcul avec structures quelconques à droite, le Tiers exclu est immédiatement dérivable, pourvu qu'on dispose de la règle de Contraction à droite :

$$\cfrac{\cfrac{\cfrac{\cfrac{A \vdash A}{0 \vdash \neg A\mathbf{;}\ A}}{0 \vdash \neg A \vee A\mathbf{;}\ A}\ \vdash\vee}{0 \vdash \neg A \vee A\mathbf{;}\ \neg A \vee A}\ \vdash\vee}{0 \vdash \neg A \vee A}\ \text{Contraction}$$

qui n'est bien sûr pas dérivable en logique intuitionniste en raison de la restriction sur les structures à droite (une seule formule). Il l'est, souvenons-nous, en logique pertinente, où Contraction est admise.

La règle d'Introduction à gauche est la suivante :

$$\frac{X(A) \vdash C \qquad X(B) \vdash C}{X\,(A \vee B) \vdash C}$$

Si Affaiblissement est autorisé, le Syllogisme disjonctif devient dérivable :

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{A \vdash A;B}}{A;\neg A \vdash B}}{\neg A; A \vdash B} \qquad B \vdash B}{\neg A; A \vee B \vdash B}$$

Affaiblissement à droite

Permutation

Remarquons qu'au passage *Ex Falso Quodlibet* a été dérivé sous la forme $A;\neg A \vdash B$

Cela ne nous dit pas encore comment la disjonction est liée au signe de ponctuation sur les structures. Il est bon de se souvenir ici de la remarque de Gentzen, selon laquelle une collection de formules, qui a la signification intuitive d'une conjonction dans l'antécédent, a la signification d'une disjonction· dans le conséquent. Dans la ligne de cette remarque, on introduit dans le langage un nouveau connecteur $\oplus$ (« fission », ou disjonction intensionnelle), qui traduit le comportement de·**;** dans les conséquents. Sa règle d'Introduction à droite est :

$$\frac{X \vdash Y\,(A;B)}{X \vdash Y\,(A \oplus B)}$$

Sa règle d'Introduction à gauche est :

$$\frac{X;A \vdash Y \qquad B;X' \vdash .Y'}{(X;A \oplus B);X' \vdash Y;Y'}$$

On a donc par la seconde règle :

$$\frac{A \vdash A \qquad B \vdash B}{A \oplus B \vdash A;B}$$

qui montre que la structure A**;**B est déductible de la formule $A \oplus B$. La première règle, elle, montre que si A**;**B a été déduit, on peut la remplacer par la formule $A \oplus B$. En fait, la disjonction intensionnelle est à l'égard de la disjonction extensionnelle comme **o** à l'égard de $\wedge$: si Affaiblissement et Contraction sont admises, les deux connecteurs se confondent. En effet, si on a $X \vdash A$, par Affaiblissement on a :

$$\dfrac{\dfrac{X \vdash A}{X \vdash A;B}}{X \vdash A \oplus B}$$

comme on a : $X \vdash A \vee B$.

De même, dans la règle d'Introduction à gauche, Affaiblissement et Contraction permettent d'identifier les contextes (comme c'était le cas avec **o**), et la règle pour $\oplus$ se confond avec la règle pour $\vee$. Les propriétés de la disjonction sont là encore héritées des propriétés qu'on accorde aux signes de ponctuations sur les structures.

Prenons comme exemple le système **LR** avec structures à droite, caractérisé par l'admission des règles structurales Associativité, Permutation, Contraction, mais sans

Affaiblissement pour **;**, avec 0 et les connecteurs introduits jusqu'ici. On peut montrer qu'il est déductivement équivalent au système **R** complet, dont le langage de base contient également **o**, la constante ***t***, où les axiomes pour ***t*** (quels qu'ils soient) assurent que A et $\boldsymbol{t} \to$ A sont interdéductibles. La constante du faux ***f*** peut être introduite comme abréviation de $\neg\boldsymbol{t}$. Il est usuel d'introduire par définition le connecteur $\oplus$, sur la base des remarques suivantes :

on a : $A\mathbf{;}B \vdash C$
<u>ssi</u> $A \vdash B \to C$
<u>ssi</u> $A \vdash \neg B\mathbf{;}C$
<u>ssi</u> $A \vdash \neg B \oplus C$

D'où la définition :

$$B \oplus C =_{\mathrm{Df}} \neg B \to C.$$

On définit récursivement une application T des structures vers les formules de la manière suivante :

$$T(A) \quad = \quad \text{A pour toute formule A}$$

$$T(0) \ = \begin{cases} \boldsymbol{t} \text{ dans l'antécédent} \\ \boldsymbol{f} \text{ dans le conséquent} \end{cases}$$

$$T(X\mathbf{;}Y) = \begin{cases} \text{dans l'antécédent, } T(X)\ \mathbf{o}\ T(Y) \\ \text{dans le conséquent, } T(X) \oplus T(Y) \end{cases}$$

$$T(X\mathbf{,}Y) = \begin{cases} \text{dans l'antécédent, } T(X) \wedge T(Y) \\ \text{dans le conséquent, } T(X) \vee T(Y) \end{cases}$$

La traduction T est étendue aux séquents de la manière suivante :

$$T(X \vdash Y) = T(X) \to T(Y).$$

PROPOSITION. — Si A est un théorème de **R,** le séquent $0 \vdash A$ est dérivable dans **LR**.

1) Si A est un axiome, on vérifie cas par cas que $0 \vdash A$ est dérivable dans **LR.** On vérifie aisément pour les axiomes de la constante ***t***, que $0 \vdash A \to (\boldsymbol{t} \to A)$, et $0 \vdash (\boldsymbol{t} \to A) \to A$ sont dérivables. Comme il a été dit plus haut dans le contexte de la Déduction naturelle, Distributivité dépend de la présence du signe de ponctuation **,**.

2) Les deux règles de **R** sont le Modus Ponens et Adjonction, et leur effet peut être obtenu. Si A est un théorème, et $A \to B$ également, par hypothèse on a $0 \vdash A$ et $0 \vdash A \to B$ prouvables dans **LR.** On a donc :

$$\cfrac{\cfrac{0 \vdash A \to B \qquad 0 \vdash A}{0 \vdash (A \to B)\ \mathbf{o}\ A}\ \mathbf{o}\text{-à droite} \qquad (A \to B)\ \mathbf{o}\ A \vdash B}{0 \vdash B}\ \text{Coupure}$$

Et il est clair que $\wedge$Introduction à droite donne les effets de la règle d'Adjonction.

COROLLAIRE. — Si $T(X \vdash Y)$ est un théorème de **R**, le séquent $X \vdash T(Y)$ est dérivable dans **LR**.

Supposons que $T(X \vdash Y)$, *i.e.* $T(X) \rightarrow T(Y)$, soit un théorème de **R** ; par la proposition précédente, $0 \vdash T(X) \rightarrow T(Y)$ est un séquent prouvable. Par ailleurs, on peut s'assurer par induction sur la complexité de X, que $X \vdash T(X)$, pour toute structure X (exemple : $0 \vdash \boldsymbol{t}$ est un axiome de **LR).** On a donc :

$$\frac{X \vdash T(X) \qquad 0 \vdash T(X) \rightarrow T(Y)}{X \vdash T(Y)}$$

Cela montre que si **R** peut prouver une implication, le conséquent est déductible (dans **LR)** de la structure dont l'antécédent du théorème est la traduction.

RÉCIPROQUE. — Si $X \vdash Y$ est dérivable dans **LR**, $T(X \vdash Y)$ est un théorème de **R.**

Si le séquent est un axiome, sa traduction évidemment est un théorème (l'axiome $0 \vdash \boldsymbol{t}$ donne $\boldsymbol{t} \rightarrow \boldsymbol{t}$). Pour les séquents dérivés, on doit montrer que si la traduction des prémisses est un théorème, la traduction de la conclusion est aussi un théorème. Quelques cas sont examinés ci-dessous :

– Le séquent est de la forme $X;\neg A \vdash Y$, et par hypothèse $T(X) \rightarrow T(A;Y)$, *i.e.* $T(X) \rightarrow A \oplus T(Y)$ est un théorème. En vertu de la définition de Ee, on a $T(X) \rightarrow (\neg A \rightarrow T(Y))$, et donc $\big(T(X)\ \mathbf{o}\ \neg A\big) \rightarrow T(Y)$, *i.e.* $T(X;\neg A) \rightarrow T(Y)$, est un théorème (voir l'interdéductibilité de $(A\ \mathbf{o}\ B) \rightarrow C$ et $A \rightarrow (B \rightarrow C)$. Le raisonnement est analogue pour l'autre règle d'introduction de la négation.

– Le séquent est de la forme $X \vdash A \rightarrow B;Y$, et par hypothèse la prémisse dont il provient a pour traduction un théorème, *i.e.* $T(X)\ \mathbf{o}\ A \rightarrow B \oplus T(Y)$ est un théorème, d'où $T(X)\ \mathbf{o}\ A \rightarrow \big(\neg B \rightarrow T(Y)\big)$ aussi. On a donc $T(X)\ \mathbf{o}\ A\ \mathbf{o}\ \neg B \rightarrow T(Y)$, d'où $T(X)\ \mathbf{o}\ \neg(A \rightarrow B) \rightarrow T(Y)$, et $T(X) \rightarrow \big(\neg(A \rightarrow B) \rightarrow T(Y)\big)$. Finalement, $T(X) \rightarrow ((A \rightarrow B) \oplus T(Y)\big)$, *i.e.* $T(X) \rightarrow T(A \rightarrow B;Y)$ est également un théorème de **R**.

– Les choses sont plus compliquées pour les règles qui non seulement substituent des formules plus complexes à des formules dans les structures (comme $\wedge$Introduction à gauche, ou $\vee$Introduction à droite), mais forment des formules plus complexes à partir de formules figurant dans des structures différentes, comme c'est le cas pour $\rightarrow$Introduction et $\oplus$Introduction à gauche. Un exemple simple aidera à suivre le raisonnement.

Supposons que le séquent dérivé soit $B \rightarrow C;A \vdash D$; il provient des deux séquents prémisses $A \vdash B$ et $C \vdash D$. Supposons que leurs traductions respectives $A \rightarrow B$ et $C \rightarrow D$ soient des théorèmes de **R** ; $(B \rightarrow C) \rightarrow (A \rightarrow D)$ l'est également, d'où $\big((B \rightarrow C)\ \mathbf{o}\ A) \rightarrow D\big)$, *i.e.* la traduction du séquent de départ. Il faut montrer que c'est toujours le cas lorsque l'implication combine des formules figurant à l'intérieur de structures quelconques.

Considérons à présent la formulation générale de la règle $\rightarrow$Introduction :

$$\frac{X \vdash Y(A) \qquad X'(B) \vdash Y'}{X'(X;A \rightarrow B) \vdash Y(Y')} \quad \rightarrow\vdash$$

En toute rigueur, il faut montrer en raisonnant par récurrence sur la complexité des traductions de X' et de Y que si les traductions des séquents prémisses sont des théorèmes, la traduction du séquent conclusion l'est aussi. Le cas simple est analogue à celui évoqué plus haut, où $Y(A) = A$ et $X'(B) = B$, X et Y' pouvant être de complexité quelconque.

Pour donner un exemple : supposons à présent que $Y(A)$ soit de la forme $F;Z(A)$; sa traduction est donc $F \oplus T\ (Z(A))$. Par hypothèse, $T(X) \to F \oplus T\ (Z(A))$ est un théorème de **R.** Par hypothèse de récurrence, sous la supposition que $T(X'(B)) \to T\ (Y')$, $T(X) \to T\ (Z(A))$ entraîne :

$$T(X'(X;A \to B)) \to T\ (Z(Y'))$$

qui donc est un théorème si $T(X'(B)) \to T\ (Y')$ l'est. Mais $T(X) \to F \oplus T\ (Z(A))$, *i.e.* $T(X) \to (\neg F \to T\ (Z(A)))$, donne par Permutation $\neg F \to (T(X) \to T(Z(A)))$, d'où par Transitivité :

$$\neg F \to [T(X'(X;A \to B)) \to T\ (Z(Y'))].$$

D'où par une nouvelle Permutation :

$$T(X'(X;A \to B)) \to (\neg F \to T\ (Z(Y'))),$$

i.e. $\quad T(X'(X;A \to B)) \to F \oplus T\ (Z(Y')),$

i.e. $\quad T\ (X'(X;A \to B)) \to T\ (Y(Y')),$

qui est donc bien un théorème si les deux traductions des prémisses du séquent le sont. Le même type de raisonnement par récurrence sur la complexité de X' où figure B donne enfin le résultat recherché. Les autres cas sont laissés au lecteur en exercice.

8. LA NÉGATION

> « Actuellement, la négation est un "sujet chaud", et il y a un besoin urgent d'une explication compréhensive de ce concept clef »
>
> Gabbay et Wansing, 1999.

Le problème de la négation a été rencontré à plusieurs reprises dans les chapitres précédents, en particulier sous la forme de la question : quels sont les rapports entre la négation dite « pertinente », la négation de De Morgan, et la négation classique ? S'agit-il vraiment de deux négations distinctes, ou bien est-ce la même négation plongée dans des contextes inférentiels différents ? Et dans la première hypothèse, qu'ont-elles en commun qui ferait d'elles deux *négations* ? Quel est le noyau central de la négation ? Le but de ce chapitre est de présenter quelques réponses plausibles à ces questions, puis d'examiner dans quelle mesure le point de vue structural est susceptible de projeter quelque lumière sur l'analyse de la négation.

LE RÉALISME DE LA NÉGATION

Certains auteurs pensent que la négation est une notion, un objet de pensée, plus précisément un opérateur portant sur des significations, ayant une forme de réalité « en soi », et que la tâche d'un système formel est de refléter aussi bien que possible ses propriétés logiques objectives. Graham Priest est sans doute l'un des représentants les plus radicaux de ce point de vue, dans la mesure où il maintient en même temps qu'il y a un et un seul objet, *la* négation, dont il s'agit de trouver la bonne théorie. C'est pourquoi il récuse le point de vue pluraliste, selon lequel il y a autant de négations, simplement liées par un air de famille, que de systèmes formels :

> « Comment donc la négation se comporte-t-elle ? Il y a une manière facile de régler cette question. Il n'y a rien de tel que la négation ; il y a plein de négations différentes : la négation booléenne, la négation intuitionniste, la négation de De Morgan. Chacune se comporte selon des règles (inférentielles, ou sémantiques) ; chacune est parfaitement légitime, et nous sommes libres d'utiliser la notion que nous voulons, aussi longtemps que nous sommes au clair sur ce que nous faisons. Si cela est correct, il n'y a rien à dire de plus sur la question, sauf ce qui justifie qu'on inscrive un connecteur dans la famille des négations. (...)
> « Je ne pense pas que cette réponse soit la bonne. (...) L'objet théorique doit s'ajuster à l'objet réel, et comment ce dernier se comporte n'est pas matière à choix » (Priest, 1999).

Mais on peut être à la fois réaliste quant à la négation, et subtilement pluraliste. La distinction entre négation déterminable, et négations déterminées, de Richard Sylvan (Routley) - l'auteur de la conception sans doute la plus sophistiquée de la négation en soi - permet de combiner les deux positions. Voici une expression claire de ce réalisme littéralement platonicien :

> « Les déterminables sont comme des formes platoniciennes ; les déterminables logiques sont des universaux d'un certain genre, auxquels ressemblent les déterminés, un peu comme les exemples ressemblent aux formes. (...)
> « On peut prétendre à bon droit que les négations sont, comme les nombres, des items de plein droit qui n'existent pas (il n'y a pas de problème d'existence pour de tels universaux). (...) La négation elle-même est un authentique non-existant, un objet qu'on ne peut croiser dans un environnement ordinaire ou extraordinaire » (Sylvan, 1999).

La distinction entre la négation comme déterminable (ou seulement partiellement déterminée, comme la négation de l'assertion par opposition, par exemple, à la négation portant sur les propriétés : honnête/malhonnête, etc.), et les négations comme déterminées, permet à Sylvan de soutenir qu'alors même que les négations particulières ont des propriétés qui les spécifient, la négation en tant que déterminable n'a pas de propriété logique qui la caractériserait. La négation est insaisissable : « *logical elusiveness of the determinable negation* ». Ce qui répond à la question de *l'unité* de la négation de système à système, et permet d'éviter l'objection d'avoir « changé de sujet » : un connecteur auquel on a retiré telle ou telle propriété usuelle de la négation exprime peut-être un autre déterminé, mais reste bien une expression de *la négation* comme déterminable. C'est bien toujours de la négation qu'il s'agit à travers différents choix de propriétés, même s'il s'agit de négations différentes, parce que la négation est à la fois une et multiple, et que la négation comme une ne semble pas avoir de propriétés véritablement essentielles (sur ce point, Sylvan se démarque d'auteurs comme Read ou Priest, pour qui le terme « la négation » est un désignateur rigide, dont la référence possède des propriétés essentielles qu'on doit pouvoir isoler) :

> « Pour avancer une thèse discutable, ni le déterminable porteur de l'unité, ni sa forme subordonnée concernant les assertions n'ont de propriétés logiques distinctives. (...)
> « Les lois traditionnelles de la pensée [*i.e.* le Tiers exclu et le Principe de non-contradiction] (...) ne valent pas en dehors des constructions traditionnelles. Elles ne valent pas, respectivement, pour les systèmes de type intuitionniste et les auteurs qui admettent l'incomplétude, et pour les systèmes du type dual de l'intuitionnisme comme les logiques paraconsistantes, et les auteurs qui admettent qu'une inconsistance explicite puisse figurer dans leurs pensées, comme c'est le cas pour certains. Une fois de plus, si l'on tente de trouver un principe commun, quelque chose à quoi adhèrent tous les penseurs, on n'obtient rien du tout, aucun principe. Aucune loi (de la négation) ne vaut universellement. (...) Bien sûr, ces positions hérétiques sont contestées. Elles le sont par l'idée que les foncteurs concernés, par exemple ceux mis en avant en abandonnant les "lois" traditionnelles, ne sont pas, ne sont pas réellement, des négations. Mais elles en sont ... » (*ibid.*).

Néanmoins, malgré ce caractère non spécifiable de la négation comme déterminable, il semble qu'on puisse en dire quelque chose, de nécessairement vague et indéterminé, naturellement. Sylvan propose ainsi de considérer les traits généraux d'opposition, de conflit, et de *polarité* comme caractéristiques typiques de la négation. Opposition ne veut pas dire nécessairement relation de contradiction : des contraires comme « un homme est honnête », « un homme est malhonnête » expriment une opposition ; un prédicat négatif comme « inélégant » ne désigne pas le complémentaire ensembliste d'« élégant », mais est borné à une sphère restreinte d'objets, etc. Aussi imprécis que soient ces traits, ils doivent être représentés dans une sémantique formelle (une modélisation) adéquate. L'idée générale est que, puisqu'il y a *conflit* entre A et ~A, un monde compatible avec un monde où ~A est vrai ne peut rendre A vrai aussi. D'où découle assez naturellement le genre de clause qu'on a vue dans le cadre de la sémantique pour **R**, mais qui peut valoir pour différentes négations déterminées :

I (~A, *a*) = 1 ssi pour tout monde *b* compatible avec *a*, **I** (A, *b*) = 0,

qui, si la relation de compatibilité est fonctionnelle, prendra la forme bien connue :

I (~A, *a*) = 1 ssi **I** (A, *a**) = 0.

Sylvan commente ainsi l'allure de ces clauses :

> « Deux traits fondamentaux de la négation normale pour l'assertion, traits qui se transposent de manière analogue aux autres formes catégorielles, sont manifestés par ces règles d'évaluation : *caractère relationnel,* passage à une situation connexe, et *polarité* explicite, inversion de la valeur assignée à cette situation » (*ibid*).

Cette caractérisation étant supposée fondamentale, la négation classique apparaît pour ce qu'elle est : un cas dégénéré, éliminant le caractère relationnel, la considération d'autres situations, au profit de la seule polarité dans le seul monde qui compte, le monde réel, le nôtre. Le problème est qu'on ne voit pas bien comment tirer, des traits de la négation dégagés jusqu'ici, une justification rationnelle de l'idée que l'évaluation de la négation *quelque part* suppose la considération *d'autres* situations, ailleurs[1]. Rien, dans l'idée d'opposition, ne motive cet aspect de relation à d'autres mondes ou situations, la relation d'opposition entre un énoncé et sa négation suggérant plutôt que deux états de chose incompatibles, ou opposés, ne peuvent avoir lieu *ensemble,* au même endroit. Il me semble qu'ici, on accorde tout simplement à la négation en soi des traits qui permettront ensuite d'obtenir les propriétés désirées : éviter que A et ~A soient nécessairement exclusifs l'un de l'autre. Mais bien sûr il y a d'autres moyens d'obtenir le même résultat, qui ne mobilisent pas des contextes ou des points connexes : une sémantique à quatre (ou trois, si l'on veut sauver le Tiers exclu en excluant ni-Vrai-ni-Faux) valeurs de vérité. Il y a donc certainement une part d'arbitraire dans cette « déduction », à partir de la négation en soi, d'une représentation de la négation qui permettra ultérieurement de prétendre que la négation pertinente est un bon déterminant : celle qui peut « tenir lieu, pour tous les buts pratiques, de la négation comme déterminable » (*ibid.*).

De plus, cette notion de la négation en tant que déterminable soulève à nouveau le problème de l'imputation d'inconsistance. Vu la manière très générale dont Sylvan la conçoit, comme opposition sous une forme ou une autre, il n'y a rien d'invraisemblable ou de contre-intuitif dans l'idée que A et ~A puissent être vrais ensembles en quelque état de choses. La négation est assez indéterminée pour recouvrir la contrariété, et on peut accorder que A et ~A peuvent être vrais ensemble, ~A étant bien la, ou une, négation de A : simplement, il s'agirait de subcontraires (comme : quelques hommes sont intelligents, quelques hommes ne sont pas intelligents). Mais alors quelle raison a-t-on de dire qu'une situation où A et ~A sont vrais ensemble est *ipso facto* inconsistante ? (C'est la substance de l'argument de Slater, 1995).

1. Le même argument se trouve sous la plume de Bull (Bull, 1986), cité par Brady dans le Prologue de Brady, 2003.

Pour souligner le caractère décevant de ce genre de déduction, il faut noter que, selon les auteurs, la légitimité généralement reconnue de telle ou telle loi de la négation peut procéder de théories très différentes de ce qu'est la négation en soi. Par exemple, pour justifier la loi de Double Négation dans le sens : $A \to {\sim}{\sim}A$, Restall invoque le caractère « naturellement » symétrique de la relation de compatibilité C entre situations qui figure dans la clause d'évaluation de la négation pertinente :

> « Considérons quelques propriétés qu'il est plausible d'attribuer à C. Par exemple, il est certain que la compatibilité *semble* être symétrique. C'est-à-dire, si xCy, alors yCx. Étant donné que tel est le cas, il s'ensuit que chaque fois que $x \vDash A$, $x \vDash {\sim}{\sim}A$ » (Restall, 1999).

Supposons en effet que $x \vDash A$, mais $x \nvDash {\sim}{\sim}A$; par la clause pour la négation en termes de situations compatibles, il existe y tel que x C y et $y \vDash {\sim}A$. Par symétrie de la relation de compatibilité, on a y C x, et donc $x \nvDash A$. Contradiction. Mais il faut noter au passage que, s'il doit sembler plausible que la relation de compatibilité soit symétrique, il *ne doit pas sembler* plausible qu'elle soit réflexive : car évidemment si $x \vDash {\sim}A$, et si x lui-même figure toujours parmi les situations compatibles avec x, on n'aura jamais $x \vDash A$ en même temps, et la possibilité d'inconsistance que la construction est destinée à assurer disparaît. Mais la non-réflexivité de la relation de compatibilité paraît purement *ad hoc*, et est aussi peu intrinsèquement plausible que l'inaccessibilité d'un monde à partir de lui-même en logique modale.

Pour Sylvan, c'est plutôt dans la mesure où la négation comme déterminable recouvre des aspects variés, dont le déterminé qu'est la « négation-sanction » : ~A compris comme : « A entraîne quelque condamnation », une absurdité par exemple, que $A \to {\sim}{\sim}A$ exprime une détermination légitime, naturelle, de la négation. Interprétée à la manière intuitionniste comme $A \to \bot$, cette négation-sanction délivre $A \to {\sim}{\sim}A$ par le seul principe d'Assertion, $A \to \big((A \to \bot) \to \bot\big)$. La loi de Double Négation dans ce sens, d'une part, spécifie une négation déterminée, d'autre part, exprime une détermination naturelle. Pour la loi de Double Négation dans l'autre sens, Sylvan considère qu'elle est plutôt consonante avec un autre aspect de la négation comme déterminable, la polarité, et l'opération qui, partant d'un pôle, fait revenir au point de départ après deux applications (négation involutive). C'est cet aspect qui est susceptible de justifier la clause selon laquelle dans la sémantique relationnelle :

$$x^{**} = x.$$

Car alors, si $x \vDash {\sim}{\sim}A$, $x^* \nvDash {\sim}A$, *i.e* . $.x^{**} \vDash A$, et donc $x \vDash A$. Sylvan conclut, en faveur de l'admission des deux sens de la loi de Double Négation :

> « Il est clair que l'option (...) Double négation au complet découle directement des explications de la négation comme polarité, en termes de notions polaires comme assertion et rejet, bon et mauvais, dire-oui et dire-non, vérité et fausseté, et autres semblables » (Sylvan 1999).

La distinction négation déterminable / négations déterminées est certainement subtile. Elle permet, on l'a dit, de maintenir qu'il s'agit toujours de la négation à travers l'attribution de propriétés différentes, ou à travers diverses représentations par des connecteurs analogues dans des systèmes différents. Elle permet un pluralisme motivé, qui admet même la négation dite « classique » dans des contextes restreints (de même, Sylvan tient que *Ex Falso*, comme le Syllogisme disjonctif, sont corrects dans une classe restreinte de situations). Mais en dépit du pluralisme originel, l'analyse de la négation est finalement censée conduire à l'idée que le meilleur déterminé est la négation pertinente, un déterminé naturel, utile, adéquat, et qu'une sémantique formelle peut convenablement traiter.

QUELQUES MOTS SUR LE DIALÉTHÉISME

Graham Priest, au contraire, défend une notion beaucoup plus déterminée de ce qu'est la vraie négation. Il remarque que nous savons distinguer les cas où l'insertion de « ne ... pas » (ou de « *not* ») permet d'obtenir une authentique négation (comme dans « Socrate est mortel / Socrate n'est pas mortel »), des cas où la véritable négation est obtenue autrement : « Quelque homme est mortel / aucun homme n'est mortel. » Il en tire la conclusion que nous avons une conception précise de la négation, indépendante des usages linguistiques de « ne ... pas ». La notion que nous saisissons sous le nom de « négation » est celle de *contradictoires* :

> « Ces exemples montrent que nous sommes capables de saisir la négation d'une manière indépendante de la façon dont "ne ... pas" fonctionne, et que nous sommes capables d'utiliser cette saisie pour déterminer quand "préfixer ne ... pas" (*notting*) nie réellement. Mais alors qu'est-ce donc que nous saisissons ? Nous voyons qu'il semble y avoir une relation d'un certain genre entre des paires comme "Socrate est mortel" et "Socrate n'est pas mortel" ; et comme "quelque homme est mortel", et "aucun homme n'est mortel". La manière traditionnelle d'exprimer cette relation est de dire que les paires sont des contradictoires, et ainsi nous pouvons dire que la relation est celle de contradiction. Les théories de la négation sont des théories au sujet de cette relation » (Priest, 1999).

Priest déduit immédiatement de cette notion de contradictoires la loi du Tiers exclu et le Principe de non-contradiction. Par définition, sont des contradictoires deux énoncés dont la disjonction est logiquement vraie, la conjonction logiquement fausse. Il en résulte en particulier que, pour tout A, $\neg$ (A $\wedge$ $\neg$A) (je respecte les choix graphiques de Priest pour le connecteur de négation, auxquels il ne faut pas accorder d'importance particulière ici). C'est ici que l'originalité de la position de Priest se manifeste le plus clairement. En maintenant le Principe de non-contradiction : « on n'a pas A et $\neg$A à la fois », qui découle de la notion même de contradictoires, Priest s'assure que $\neg$A exprime bien le contradictoire de A. Par là, il est répondu à l'objection évoquée plus haut, selon laquelle il se pourrait que, en attribuant certaines propriétés à $\neg$A, on cesse du même coup de parler du véritable contradictoire de A. Et même en cas de vérité simultanée de A et $\neg$A, c'est donc bien de la vérité d'un énoncé et de son contradictoire dont il sera question. Pour le TND, Priest ajoute l'argument suivant : ·

> « Un authentique opérateur formant un contradictoire sera un opérateur, qui appliqué à un énoncé A, couvre *tous* les cas où A n'est pas vrai. C'est donc un opérateur $\neg$ tel que $\neg$A est vrai ssi A est soit faux soit ni vrai ni faux. (En français (*English*), un tel opérateur pourrait être : ce n'est pas le cas que.) Pour cette notion, qui est le réel opérateur formant le contradictoire, le Principe du Tiers exclu vaut » (*ibid.*).

Un peu plus loin dans le même article, Priest propose de définir la fausseté de la manière suivante :

A est faux signifie par définition : $\neg$A est vrai.

Vu la nécessité de couvrir tous les cas, la fausseté de A telle que définie couvre aussi bien le cas où A est faux au sens usuel, que le cas où A ne serait ni vrai ni faux, ce qui peut s'accorder avec l'une des sémantiques formelles proposées par Priest, où il n'y a que deux valeurs, Vrai et Faux, mais où il est admis que tout énoncé est dans une certaine relation soit avec l'un, soit avec l'autre de ces objets ; ce qui valide le Tiers exclu[1]. Quoi qu'il en soit ; les lois de Double Négation sont assurées par le caractère symétrique de la relation de contradiction : ¬A étant le contradictoire de A, A l'est aussi de ¬A, et est donc identique au contradictoire ¬¬A de ¬A. Les lois de De Morgan découlent de la définition de la fausseté et des lois de Double Négation. Il est à nouveau frappant de constater comment des considérations tout à fait divergentes sur la négation « en soi » permettent finalement de justifier, à partir de points de départ différents, les mêmes lois.

Mais le cœur du « dialéthéisme » n'a pas encore été atteint par ces remarques. La thèse qui le caractérise est que non seulement un même énoncé peut être, par exemple, dans des situations impossibles, à la fois vrai et faux, mais qu'il y a effectivement, dans notre monde, des énoncés qui sont à la fois vrais et faux. Sous la définition de la fausseté donnée plus haut, un tel énoncé A étant faux, sa négation ¬A est vraie ; mais A étant également vrai, ¬A est faux par ailleurs. Il en résulte que A et ¬A peuvent être vrais tous les deux, et qu'il y a des contradictions vraies.

> « Il est commun de distinguer entre la Loi du Tiers exclu et le Principe de Bivalence : toute affirmation est soit vraie soit fausse. Bien que ce soient des compagnons naturels, l'un peut valoir sans l'autre, étant donné une explication correcte d'autres choses. De même, nous devons distinguer entre la Loi de non-contradiction et ce que j'appelle, faute d'un meilleur terme, le Principe de Consistance : aucune affirmation n'est à la fois vraie et fausse. À nouveau, bien que ce soient des compagnons naturels, il est tout à fait possible d'avoir l'un sans avoir l'autre. En particulier, comme je l'ai déjà fait remarquer, le fait que toute instance de ¬ (A ∧ ¬A) soit vraie n'empêche pas, à elle seule, que des instances de A et ¬A soient vraies. Que faut-il donc dire du Principe de Consistance ? » (*ibid.*).

Il faut l'abandonner, et vérité et fausseté peuvent donc être attribuées en même temps. Il faut néanmoins observer que la possibilité que A et ¬A soient vrais ensemble, fondée sur la possibilité que A soit à la fois vrai et faux, n'est pas tirée à proprement parler d'une analyse de la négation, mais d'arguments concernant les paradoxes sémantiques (je ne reprendrai pas ici l'analyse de Priest ; disons simplement ici qu'il faut accepter comme un fait que l'énoncé A qui dit « A n'est pas vrai » est à la fois vrai et faux). L'abandon du Principe de Consistance, qui fait que la conjonction A ∧ ¬A peut être vraie, semble en contradiction manifeste avec l'analyse de la notion de contradictoires, qui justifie ¬ (A ∧ ¬A). Priest répond à cette objection en faisant remarquer que si le dialéthéisme est vrai, il n'y a aucun problème à ce que à la fois ¬(A ∧ ¬A) et (A ∧ ¬A) soient vrais, du moins pour quelque énoncé A, puisque des contradictions peuvent être vraies. En fait, « vrai » a perdu ici sa signification ordinaire, qui est confinée dans la valeur « seulement vrai », si on opère avec une sémantique à trois valeurs : si A est à la fois vrai et faux - paradoxal - ¬(A ∧ ¬A) peut bien être vrai selon les tables de vérité pour la négation et la conjonction, mais c'est au sens de « paradoxal » ; et il en est de

1. Dans cet article, Priest utilise des relations entre les formules et les deux valeurs de vérité classiques, plutôt que des valuations dans un ensemble à trois valeurs de vérité, Vrai, Faux, Paradoxal, *i.e.* à la fois vrai et faux (Priest, 1979).

même pour (A ∧ ¬A). Donc le Principe de non-contradiction est certes vrai, mais il est plus exactement paradoxal. Peut-être n'y a-t-il là aucun problème pour le dialéthéiste, mais il est difficile de voir quel est le contenu ou le statut d'un « principe », censé exprimer une propriété fondamentale de la négation, qui certes est toujours vrai, mais est aussi faux. La thèse dialéthéiste est-elle, elle aussi, à la fois vraie et fausse (paradoxale au sens usuel, elle l'est assurément) ?

Comme on peut s'y attendre, la thèse dialéthéiste invalide aussi bien *Ex Falso Quodlibet* que le Syllogisme disjonctif, qui ne sont donc pas des « lois » de la négation. C'est pourquoi je crains qu'il n'y ait quelque futilité dans ces tentatives pour découvrir la nature de la négation « en soi », et en tirer de quoi justifier les propriétés jugées désirables : car des conceptions très différentes aboutissent finalement, peu ou prou, à valider les mêmes résultats, dès lors qu'on les désire, et à invalider les mêmes inférences, dès lors qu'on les rejette.

LA NÉGATION EN TERMES DE RÔLE INFÉRENTIEL

Les différents systèmes en Calcul des séquents à conclusion multiple offrent une base très différente pour l'analyse d'un connecteur de négation. Je laisserai de côté ici la question de savoir si l'analyse du rôle inférentiel d'un tel connecteur est susceptible d'épuiser le *meaning*, la signification du connecteur[1]. Parlons donc simplement d'une négation caractérisée par les deux règles (comme au chap. 7) :

$$\frac{X;A \vdash Y}{X \vdash \neg A;Y}\ (\vdash\neg) \qquad\qquad \frac{X \vdash A;Y}{X;\neg A \vdash y}\ (\neg\vdash)$$

où l'on peut, pour simplifier les choses, penser X et Y comme des multi-ensembles de formules (Associativité et Permutation sont donc présupposées). Cette représentation de la négation a-t-elle un intérêt conceptuel particulier ?

Comme c'est le cas pour les autres connecteurs, la négation est caractérisée par ces mêmes deux règles, pour tous les systèmes sur la base d'un langage qui contient ce connecteur : systèmes classiques, intuitionnistes, de logique pertinente, etc. Les différences proviennent chaque fois de la disponibilité des règles d'affaiblissement et de contraction à droite et à gauche, ainsi que de la restriction ou non des conséquents à une seule formule. Il n'est nullement suggéré, par un tel point de vue, que l'on change de négation en passant d'un système à l'autre, puisque seules les propriétés variées qu'on reconnaît à la relation de déductibilité sont en jeu et caractérisent les différents systèmes.

Une objection à laquelle on pourrait penser est que cette représentation uniforme de la négation dans différents systèmes est purement relative à ce Calcul, qui n'est qu'un système parmi d'autres, et que sous d'autres présentations (autant de manières d'engendrer la relation ⊢), par exemple en Déduction naturelle, la distinction des

1. Sur cette question, voir par exemple le débat entre :Michael Hand et Neil Tennant, *in* Gabbay-Wansing, 1999.

connecteurs de négation reprend ses droits, puisque la différence des systèmes est marquée par l'introduction de nouvelles règles. Il y aurait donc quelque chose d'inessentiel dans le Calcul des séquents à conclusion multiple, qui ne justifierait du coup aucune conclusion sur l'identité de la négation. Il me semble qu'un résultat avancé par Arnon Avron est susceptible de lever cette objection (Avron, 1999). La thèse de cet article est que « la présence d'une négation interne est ce qui rend une logique essentiellement *à conclusion multiple* »[1]. Donc, même quand un système déductif n'a pas la forme explicite d'un Calcul des séquents à conclusion multiple, on peut le considérer comme *essentiellement* tel, pourvu que la négation soit munie des règles qui font d'elle une négation interne. D'où la possibilité de considérer le Calcul à conclusion multiple comme une forme particulièrement explicite, stable quant aux règles pour la négation, de systèmes moins explicites.

Disons qu'un connecteur unaire est une négation interne pour une relation de conséquence à conclusion multiple si et seulement si cette relation est close pour les deux règles ci-dessus[2]. Soit $\vdash$ une relation de conséquence à conclusion unique : $\neg$ est une négation interne pour $\vdash$, s'il existe une relation de conséquence $\vdash^*$ à conclusion multiple telle que :

1) $X \vdash^* A$ <u>ssi</u> $X \vdash A$;
2) $\neg$ est une négation interne pour $\vdash^*$.

Une relation de conséquence à conclusion unique est *essentiellement à conclusion multiple* <u>ssi</u> elle a une négation interne. On peut démontrer la proposition suivante :

THÉORÈME. — $\neg$ est une négation interne pour une relation $\vdash$ à conclusion unique <u>ssi</u> les conditions suivantes sont réunies :

(i) $A \vdash \neg\neg A$
(ii) $\neg\neg A \vdash A$
(iii) Si $X, A \vdash B$, alors $X, \neg B \vdash \neg A$.

Autrement dit, tout système qui permet de dériver les lois de Double Négation et Contraposition, est essentiellement à conclusion multiple et équivalent (en un sens restreint : pour les conclusions-formules) à un système où figurent les deux règles pour la négation mentionnées ci-dessus.

Esquisse de la preuve : il est évident en vertu de la définition, que si un système à conclusion unique a une négation interne, la relation de conséquence qu'il engendre satisfait les trois conditions. Reste à montrer qu'elles sont suffisantes. On définit une relation de déductibilité $\vdash_s$ telle que :

$$A_1, ..., A_n. \vdash_s B_1 ..., B_k$$

---

1. Avron parle de « connecteurs internes », et en particulier de « négation interne » à propos de ce qu'on appelle d'ordinaire « connecteurs intensionnels » (ou « multiplicatifs » en logique linéaire) et « négation intensionnelle ». Avron fait également la distinction (relativement à un système déductif) entre une relation de conséquence *externe*, selon laquelle une conclusion C est conséquence de prémisses s'il y a une preuve dans le système de C à partir des prémisses, et une relation de conséquence *interne*, selon laquelle C est conséquence de prémisses si l'implication emboîtée des prémisses et de la conclusion est un théorème du système. Sur cette distinction, voir ci-dessus les premières lignes du chapitre 7.

2. Ce n'est pas exactement la définition d'Avron, mais cette condition lui est équivalente.

<u>ssi</u> pour tous $i = 1, \ldots, n$, et $j = 1, ..., k$, on a à la fois :

$$A_1, ..., A_{i-1}, \neg B_i \ldots, \neg B_k, A_{i+1}, ... , A_n \vdash \neg A_i$$
$$A_1, ..., A_n, \neg B_1, \ldots , \neg B_{j-1}, \neg B_{j+1}, ... , \neg B_k \vdash B_j.$$

Il faut vérifier qu'on a bien $\Gamma \vdash_s A$ <u>ssi</u> $\Gamma \vdash A$, et que $\neg$ est bien une négation interne pour $\vdash_s$. Il est évident que si $A_1, ..., A_n. \vdash_s B$, alors la condition $A_1, ..., A_n. \vdash B$ est réalisée. Par la condition (iii) plus haut, si $A_1, ..., A_n. \vdash B$, on a $A_1, ..., A_{i-1}, \neg B \vdash \neg A_i$ pour tous les A_i donc la réciproque : si $A_1, ..., A_n. \vdash B$, alors $A_1, ..., A_n. \vdash_s B$ est vraie. Il faut enfin vérifier que $\neg$ est une négation interne pour $\vdash_s$. Si on a $\Gamma, A \vdash_s \Delta$ par la condition (ii), $\neg\neg A \vdash A$, d'où $\vdash_s A, \neg A$, d'où par Coupure : $\Gamma \vdash_s \Delta, \neg A$. De même, $\Gamma \vdash_s \Delta, A$, par la condition (i) $A \vdash \neg\neg A$, d'où $A, \neg A \vdash_s$, et par Coupure : $\Gamma, \neg A \vdash_s \Delta$.

La conclusion de ces remarques est bien évidemment que des systèmes où les règles pour la négation peuvent se présenter différemment, et où les règles structurales induisent des relations de dérivabilité distinctes, pourvu qu'ils satisfassent les trois conditions mentionnées plus haut, ont essentiellement la même négation interne, et sont du même coup essentiellement à conclusion multiple. Au nombre de ces logiques, la logique classique bien sûr, les différents systèmes de logique pertinente, la logique linéaire. En revanche, la logique intuitionniste ne possède pas une telle négation. Il semble donc que la représentation de la négation en termes de règles d'inférence justifie l'idée que la logique classique et la logique pertinente font usage de la *même* négation. Et nous aurions là la réponse au débat prolongé mené, à travers différents chapitres, à propos de l'identité ou de la pluralité des négations dans ces deux logiques.

LES DEUX POINTS DE VUE

Le problème vient du fait qu'on peut caractériser autrement la négation, en termes de modèles, et que la « négation sémantique forte », liée aux Principes du Tiers exclu et de non contradiction, n'apparaît plus que comme une négation parmi d'autres. L'idée d'Avron est la suivante. On peut identifier un modèle à une théorie, l'ensemble des énoncés vrais dans le modèle au sens usuel du terme, quelque chose comme une description d'état au sens de Carnap. Un modèle est ordinairement conçu comme décidant tout énoncé du langage (Principe du Tiers exclu), et ne vérifiant pas à la fois un énoncé et sa négation (Principe de non-contradiction). Sous l'identification d'un modèle à une théorie, cela revient à ne considérer que des théories consistantes et complètes. Un connecteur unaire d'un langage exprime la négation sémantique forte, s'il « reflète ce point de vue », c'est-à-dire si la sémantique le dote de propriétés telles que $\neg A$ est vrai dans un modèle exactement quand A est faux. L'admission de théories (modèles) inconsistantes et / ou incomplètes dans une sémantique peut donc permettre de mesurer le degré selon lequel une logique contient une négation sémantique, en ce sens précisé. Si par exemple la relation de déductibilité qui caractérise telle logique coïncide avec une relation de conséquence définie en termes de théories à la fois consistantes et complètes, on pourra dire que cette logique contient la négation sémantique forte : dans la terminologie d'Avron, elle est dite « fortement normale », *i.e.* fortement complète relativement à l'ensemble des théories consistantes et complètes (c'est évidemment le cas de la logique classique).

> « La loi de non-contradiction veut dire de manière interne que seules des théories consistantes peuvent avoir un modèle, tandis que la loi du tiers exclu veut dire que l'ensemble des énoncés qui sont vrais dans un modèle doit être complet quant à la négation. Les ensembles de théories consistantes, de théories complètes et de théories normales (celles qui sont les deux à la fois), ont donc une importance cruciale quand nous voulons trouver le degré auquel un connecteur unaire donné d'une logique peut être pris comme une négation sémantique. Ainsi des théories complètes reflètent un état de choses où vaut la loi du tiers exclu. Il est donc raisonnable de dire que cette loi vaut pour une logique L si sa relation de conséquence ⊢L est *déterminée* par son ensemble de théories complètes. De même, L satisfait (fortement) la loi de non-contradiction ssi ⊢L est déterminée par son ensemble de théories consistantes, et elle satisfait sémantiquement les deux lois ssi ⊢L est déterminée par son ensemble de théories normales » (Avron, 1999).

Soit L une logique : une sémantique S pour L est un ensemble non vide de points, qui sont des théories (ensembles de formules clos pour la relation de déductibilité selon L, en excluant l'ensemble de toutes les formules). Un modèle d'une formule A est une théorie à laquelle A appartient. Un modèle d'une théorie T est n'importe quel point qui est une extension de T. Une formule A est S-valide ssi tout point de S est un modèle de A. Enfin une formule A est S-conséquence de T ssi tout modèle de T est modèle de A.

Étant donné une logique L, et les notions usuelles de théories consistantes et de théories complètes selon la relation de déductibilité de L, on peut considérer pour spécifier S l'ensemble des théories consistantes, l'ensemble des théories complètes, et celui des théories normales (qui sont les deux à la fois). On a donc trois sémantiques, qui permettent de déterminer, pour chaque logique, trois relations de conséquences distinctes : selon que les modèles sont les théories consistantes, les théories complètes, ou les théories normales. Je donne quelques exemples des résultats obtenus à partir de ces définitions[1].

Une logique L est *fortement consistante* ssi la relation de déductibilité $\vdash_L$ est équivalente à la relation de conséquence définie en termes de modèles (théories) consistants. La logique pertinente *n'est pas* fortement consistante (alors que la logique classique l'est). Soit Γ un ensemble de formules inconsistant, T la clôture déductive de Γ, et A une formule telle que A ne soit pas déductible de T selon **R** (il y en a, puisque EFQ n'est pas un théorème de **R**). Cependant il est vrai que tout modèle consistant de T est modèle de A, puisqu'il n'y a tout simplement pas d'extension consistante de T. Donc, si la relation de conséquence est définie en termes de théories consistantes, A est conséquence de T (et de Γ). Donc **R** n'est pas fortement consistant.

Une logique est *fortement c-normale, i.e.* fortement *c*-complète relativement à l'ensemble des théories normales, si pour toute théorie consistante T et toute formule A, A est déductible de T ssi A est conséquence de T au sens des modèles ·à la fois consistants et complets. **R** *n'est pas* fortement *c*-complet relativement à l'ensemble des théories normales. Supposons par exemple que T contienne $\neg A$ et $A \vee B$, et que B ne soit pas déductible de T (ce qui est possible puisque le SD n'est pas un théorème de **R**). Cependant tous les modèles consistants et complets de T contiennent B ; sinon, il y aurait un modèle qui contiendrait $\neg B$, donc $\neg A \wedge \neg B$ par Adjonction, donc $\neg (A \vee B)$

1. Je ne prétends évidemment pas résumer l'article très dense d'Avron, mais seulement mentionner quelques résultats qui me paraissent particulièrement intéressants du point de vue de ce chapitre.

par les lois de De Morgan, lequel modèle serait inconsistant. Donc B est conséquence de $\neg A$ et $A \vee B$ pour cette classe de modèles[1].

En revanche, **R** est *fortement complet* relativement à l'ensemble des théories complètes, *i.e.* la relation de déductibilité selon **R** coïncide avec la relation de conséquence définie en termes de théories complètes. Si A est déductible selon **R** d'un ensemble Γ, toutes les théories qui sont des extensions complètes de Γ contiennent A, *i.e.* A est conséquence de Γ pour la relation définie en termes de modèles complets. Inversement, si A n'est pas déductible de Γ (A n'appartient pas à la clôture déductive T de Γ), on n'a pas à la fois, pour n'importe quelle formule B, $\Gamma, B \vdash A$ et $\Gamma, \neg B \vdash A$. Sinon on aurait $\Gamma, B \vee \neg B \vdash A$, d'où $B \vee \neg B$ étant un théorème de **R,** $\Gamma \vdash A$. On peut donc étendre Γ en une théorie complète qui ne contient pas A.

R est *faiblement normal, i.e.* faiblement complet relativement à l'ensemble des théories normales, au sens où A est un théorème de **R** ssi A est valide relativement à toutes les théories complètes et consistantes (pour la relation de déductibilité de **R**)**.** Supposons que A ne soit pas un théorème ; alors $\{\neg A\}$ est consistant. Sinon, pour quelque formule B, on aurait $\neg A \vdash (B \wedge \neg B)$, d'où $A \vee \neg A \vdash A \vee (B \wedge \neg B)$, de sorte que $A \vee (B \wedge \neg B)$ serait un théorème de **R.** Comme par ailleurs $\neg (B \wedge \neg B)$ est un théorème de **R,** par la Règle γ, $\vdash A$, contrairement à l'hypothèse[2]. Donc $\{\neg A\}$ est consistant. On montre ensuite que tout ensemble consistant de formules a une extension consistante et complète (normale) ; une telle extension de $\{\neg A\}$ ne contient évidemment pas A.

Ces quelques résultats suggèrent les conclusions suivantes. La logique classique est fortement normale, *i.e.* sa relation de déductibilité coïncide avec la relation de conséquence déterminée par les théories consistantes et complètes à la fois. Sa négation est donc la négation sémantique forte, pour laquelle un modèle est toujours consistant et complet. La logique pertinente, bien qu'elle soit fortement complète (relativement aux théories complètes), n'est pas fortement consistante, *i.e.* sa relation de déductibilité ne coïncide pas avec la relation de conséquence définie à partir des théories consistantes prises comme modèles. Elle est cependant faiblement normale, au sens où les théorèmes de **R** sont valides dans les sémantiques qui ne considèrent que des modèles consistants et complets, les modèles normaux. Du point de vue sémantique, il y a donc une différence entre la négation classique, et la négation qui figure dans la logique pertinente. Ces remarques laissent évidemment une nouvelle question ouverte : de ces deux points de vue, l'un syntaxique, l'autre sémantique, quel est celui qui donne la meilleure représentation de la négation dans les logiques pertinentes ? Je pencherais personnellement pour le point de vue qui caractérise la négation en termes de rôle inférentiel, identifie la négation classique et la négation pertinente, et fait porter le poids des différences sur les règles structurales qui spécifient différentes relations de déductibilité. Mais je reconnais que le problème est difficile, et qu'aucune réponse ne

1. Les preuves détaillées de ces propositions sont données dans Avron, 2002 *a*.

2. La règle γ dit que si $\neg$ A est un théorème, et si $A \vee B$ est aussi un théorème, alors B est un théorème. C'est, si l'on veut, le syllogisme disjonctif restreint au cas particulier où les prémisses sont des théorèmes. Le caractère admissible de la règle γ a été démontré par Meyer et Dunn dans Anderson et Belnap, 1975.

semble s'imposer indiscutablement. La boîte de Pandore de la négation n'est pas prête d'être refermée.

BIBLIOGRAPHIE

Ackermann Wilhelm, « Begründung einer strengen Implikation », *Journal of Symbolic Logic,* 21, 1956.

Anderson Alan Ross et Belnap Nuel, « The pure calculus of entailment », *Journal of Symbolic Logic*, 27, 1962.

Anderson Alan Ross et Belnap Nuel, *Entailment, the Logic of Relevance and Necessity,* vol. I, Princeton UP, 1975.

Anderson Alan Ross, Belnap Nuel et Dunn Michael, *Entailment, the Logic of Relevance and Necessity,* vol. II, Princeton UP, 1992.

Aristote, *Premiers Analytiques,* d'après l'édition bilingue de la Loeb Classical Library, Harvard UP, 2002 (1938 pour la première édition).

Avron Arnon, « Whither relevance logic? »,*Journal of Philosophical Logic,* 21, 1992.

Avron Arnon, « Negation : Two points of view », *in* Gabbay et Wansing, 1999.

Avron Arnon, « On negation, completeness, and consistency », *in* Gabbay et Guentlmer, 2002 *a.*

Barwise Jon et Perry John, *Situations and Attitudes,* MIT Press, 1983.

Belnap Nuel, « Tonk, Plonk and Plink », *Analysis,* 22, 1962.

Belnap Nuel, « Display logic »,*Journal of Philosophical Logic,* 11, 1982.

Brady Ross (ed.), *Relevant Logics and their Rivals,* vol. II, Ashgate PC, 2003.

Bull Robert, « Review of *Relevant Logic and their Rivals,* vol. I », *Australasian Journal of Philosophy*, 64, 1986.

Carnap Rudolf, *Meaning and Necessity,* The University of Chicago Press, 1947, trad. fr. *Signification et· nécessité,* Paris, Gallimard, 1997.

Church Alonzo, « The weak theory of implication », *in* Menne, Wilheimy, Angsil (eds), *Kontrolliertes Denken,* Munich, 1951.

Copeland B. J., « On when a semantics is not a semantics : Some reasons for disliking the Routley-Meyer semantics for relevance logic », *Journal of Philosophical Logic*, 8, 1979.

Copeland B. J., « What is a semantics for classical negation? », *Mind*, 1986.

Curry Haskell B. et Feys R., *Combinatory Logic,* vol. I, Amsterdam, North Holland, 1958.

Došen Kosta, « A historical introduction to substructural logics », *in* Schroeder-Heister et Došen, 1993.

Dummett Michael, *Elements of Intuitionism,* Oxford UP, 1977.

Dunn Michael, « Relevance logic and entailment », *in* Gabbay et Guentlmer, 1986.

Dunn Michael, « A comparative study of various model-theoretic treatments of negation : A history of formal negation », *in* Gabbay et Wansing, 1999.

Etchemendy John, « Tarski on truth and logical consequence », *Journal of Symbolic Logic*, 53, 1988.

Etchemendy John, *The Concept of Logical Consequence*, Harvard UP, 1990.

Fine Kit, « Semantics for quantified relevance logic », *Journal of Philosophical Logic,* 17, 1988.

Fitch F. B., *Symbolic Logic*, New York, Ronald Press, 1952.

Gabbay Dov et Guenthner F. (eds), *Handbook of Philosophical Logic,* l^{ère} éd., vol. 3, Dordrecht, P. C. Reidel, 1986.

Gabbay D. et Guenthner F. (eds), *Handbook of Philosophical Logic,* 2ème éd., vol. 4, Dordrecht, A. P. Kluwer, 2002.

Gabbay D. et Guenthner F. (eds), *Harulbook of Philosophical Logic,* 2ème éd., vol. 9, Dordrecht, A. P. Kluwer, 2002 *a.*

Gabbay D. et Wansing Heinrich (eds), *What is Negation?,* Dordrecht, Kluwer, 1999.

Geach Peter, « Entailment », in *Logic Matters,* Oxford, Blackwell, 1972.

Geach Peter, « Review of *Entailement* », *Philosophy,* 52, 1977.

Gentzen Gerhard, « Untersuchungen über das logische Schließen », *Mathematische Zeitschrift,* 39, 1934 ; trad. fr. « Recherches sur la déduction logique », Paris, PUF, 1955.

Gentzen Gerhard, « Die Widerspruchsfreiheit der reinen Zahlentheorie », *Mathematische Annalen 112,* 1936, trad. angl. in *The collected Papers of Gerhard Gentzen,* M. E. Szabo (ed.), Amsterdam, North Holland, 1969.

Girard Jean-Yves, « Linear logic », *Theoretical Computer Science,* 50, 1987.

Grice Paul (1967), « Logic and conversation, William James lectures », in *Studies in the Way of Words*, Harvard UP, 1989.

Haack Susan, *Philosophy of Logic,* Cambridge UP, 1978.

Heyting Arend, « Die formalen Regeln der intuitionischen Logik », *Sitzungsberichte der Preußischen Akademie der Wissenschaften,* 1930.

Heyting Arend, *Intuitionism : An Introduction,* Amsterdam, North Holland, 1956.

Hilbert David et Bernays Paul, *Grundlagen der Mathematik,* Springer, 1934 ; trad. fr. *Fondements des mathématiques,* par François Gaillard et Marcel Guillaume, Paris, L'Harmattan, 2001.

Hughes G. E. et Cresswell M. J., *An Introduction to Modal Logic,* Londres, Routledge, 1968.

Jonsson Bjarni et Tarski Alfred, « Boolean algebras with operators : Part I », *American Journal of Mathematics,* 73, 1951.

Kneale William et Kneale Martha, *The Development of Logic,* Oxford, Clarendon Press, 1962.

Kripke Saul, « The problem of entailment »,*Journal of Symbolic Logic,* 24, 1959.

Kron Alexsandar, « Deduction theorems for relevant logics », *:Zeitschrift für mathematische Logik und Grundlagen der Mathematik,* 19, 1973.

Kron Alexsandar, « Deduction Theorems for TER, reconsidered », *:Zeitschrift für mathematische Logik und Grundlagen der Mathematik,* 22, 1976.

Lewis David, *Counterfactuals,* Basil Blackwell, 1973.

Lewis C. I. et Langford C. H., *Symbolic Logic,* New York, Dover, 1932.

Mares Edwin et Fuhrmann A., « A relevant theory of conditionals », *Journal of Philosophical Logic,* 24, 1995.

Meyer Robert, Routley Richard et Dunn Michael, « Curry's paradox », *Analysis,* 39, 1979.

Meyer Robert et Sylvan Richard, « Extensional reduction, II », *in* Brady, 2003.

Myhill John, « Real implication », *in* Norman et Sylvan, 1989.

Norman Jean et Sylvan Richard (eds), *Directions in Relevant Logics,* Dordrecht, Kluwer, 1989.

Nute Donald et Cross Charles B., « Conditional logic », *in* Gabbay et Guenthner, 2002.

Parry William, « Analytic implication », *in* Norman et Sylvan, 1989.

Prawitz Dag, *Natural Deduction,* Stockholm, Almquist & Wiksell, 1965.

Priest Graham, « The logic of paradox »,*Journal of Philosophical Logic,* 8, 1979.

Priest Graham, « What not? A defence of dialetheic theory of negation », *in* Gabbay et Wansing, 1999.

Priest Graham, *An Introduction to Non-Classical Logic,* Cambridge UP, 2001.

Priest Graham, Routley Richard, Norman Jean (eds), *Paraconsistent Logic, Essays on the Inconsistent,* Munich, Philosophia Verlag, 1989.

Priest Graham et Sylvan Richard, « Simplified semantics for basic relevant logics », *Journal of Philosophical Logic,* 21, 1992.

Quine Willard V., *Mathematical Logic,* Harvard UP, 1940.

Quine Willard V., *Methods of Logic,* New York, Holt, 1950 ; trad. fr. *Méthodes de logique,* par M. Clavelin, Paris, Armand Colin, 1972.

Quine Willard V., *Elementary Logic* (éd. révisée de 1941), Harvard UP, 1966 ; trad. fr. *Logique élémentaire,* par J. Largeault, Paris, Armand Colin, 1972.

Quine Willard V., *Philosophy of Logic,* Englewood Cliffs Prentice Hall, 1970.

Quine Willard V., *Quiddities,* Harvard UP, 1987.

Read Stephen, « What is wrong with disjunctive syllogism? », *Analysis,* 41, 1981.

Read Stephen, *Relevant Logic : A Philosophical Examination of Inference,* Oxford, Blackwell, 1988.

Read Stephen, *Thinking about Logic,* Oxford UP, 1995.

Restall Greg, « Negation in relevant logics (How 1 stopped worrying and learned to love the Routley Star) », *in* Gabbay et Wansing, 1999.

Restall Greg, *An Introduction to Substructural Logics,* Londres, Routledge, 2000.

Rivenc François, *Sémantique et vérité : de Tarski à Davidson,* Paris, PUF, 1998.

Routley Richard, Meyer Robert, « Semantics of entailment », *in* Hughes Leblanc (ed.), *Truth, Syntax, and Modality,* Amsterdam, North Holland, 1973.

Routley Richard, Meyer Robert, Plumwood Val et Brady Ross, *Relevant Logics and their Rivals,* vol. I, P. C. Ridgeview, Atascadero, 1982.

Schroeder-Heister Peter et Došen Kosta (eds), *Substructural Logics,* Oxford, Clarendon Press, 1993.

Slater B. H., « Paraconsistent logics »,*Journal of Philosophical Logic,* 24, 1995.

Stone M. H., « Topological representations of distributive lattices and Brouwerian logics », *Casopis pro pestovani matematiky a fysiki*, 67, 1937.

Strawson Peter F., *Introduction to Logical Theory,* Londres, 1952.

Sundholm Göran, « Systems of deduction », *in* Gabbay et Guenthner, 1983 (vol. 1 de Gabbay et Guenthner, 1986).

Sylvan Richard (anciennement Routley), « What is that item designated negation? », *in* Gabbay et Wansing, 1999.

Urquhart Alasdair, « Semantics for relevant logics »,*Journal of Symbolic Logic,* 37, 1972.

Urquhart Alasdair, « The undecidability of entailment and relevant implication », *Journal of Symbolic Logic,* 49, 1984.

Urquhart Alasdair, « Semilattice semantics for relevance logics », Anderson et Belnap, 1992, § 47 et s.

Woodger J. H. (1956) (trad.), *Alfred Tarski : Logic, Semantics, Metamoihematics,* P. C. Hackett, 2' éd., 1983.